Arenberg Doctoral School

Faculty of Bioscience Engineering

Laboratory of Gene Technology

Faculty of Biological Sciences

Department of Pathogen Biology

and Immunology

Characterization of lytic bacteriophages infecting *Pseudomonas aeruginosa* and their peptidoglycan and exopolysaccharide degrading enzymes

Katarzyna Danis-Włodarczyk

Supervisors:

Prof. Rob Lavigne, KU Leuven, Belgium

Prof. Zuzanna Drulis-Kawa, University of Wrocław, Poland

Examination committee:

Prof. Krystyna Dąbrowska, Polish Science Academy, Poland

Prof. Andrew Kropinski, University of Guelph, Canada

Prof. Abram Aersten, KU Leuven, Belgium

Prof. Jan Michiels, KU Leuven, Belgium

Prof. Ivo Vankelecom, KU Leuven, Belgium

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Bioscience Engineering and Doctor in Biological Sciences

December 2016

For my dad

Mirosław Danis

(04.03.1954 - 24.05 1985)

Table of contents

Summary

Antibiotic resistance of bacterial pathogens is an emerging problem worldwide. As multidrug resistant organisms fail to respond to antimicrobial therapy, infections become more severe and cause more complications which in turn leads to longer illnesses (Gould, 2006; Haecker, 2009; Hübner *et al.*, 2012). The aerobic Gram-negative bacterium *Pseudomonas aeruginosa* is one of the most important and dangerous microbes inhabiting the hospital environment. It has the ability to adapt to and thrive in many ecological niches, from water and soil environments to plant and animal tissues. *P. aeruginosa* also possesses a wide array of virulence factors, that do not only cause extensive tissue damage, but also interfere with the human immune system (Dzierżanowska, 2008). As an opportunistic pathogen it is particularly known for causing endogenous infections in immune-deficient individuals (AIDS, cystic fibrosis, cancer) and presents resistance to many antibiotics (Cornelis *et al.*, 2008).

During chronic infections populations of *P. aeruginosa* undergo characteristics evolutionary adaptations, including reduction of virulence factors production and transition from planktonic form to biofilm-associated lifestyle. The biofilm play an important role in evolution of high-level antibiotic resistance and protects the embedded bacteria from antimicrobial agents applied from without. Moreover, during prolonged infections coexistence of divergent phenotypic lineages of *Pseudomonas* within patients was observed, which makes accurate diagnosis and treatment even more challenging (Winstanley *et al.*, 2016).

For these reason, new antibacterial therapies are urgently needed. With hurdles in the antibiotic research and development pipelines, especially against Gram-negative pathogens, the scientific community is exploring the development of alternative forms of antimicrobial therapies. The idea of using bacteriophages as natural parasites of bacteria is well known. Bacteriophages have co-evolved intimately with their host for three billion years, and therefore highly efficient antibacterial mechanisms have emerged, granting them unique advantages over classical antibiotics to kill bacteria. Indeed, phage cocktails are commonly applied as alternative or as supportive treatments simultaneously with antibiotics in Eastern Europe, as part of standard care. Positive results are routinely obtained with the eradication of *Escherichia, Pseudomonas, Proteus, Klebsiella* and *Staphylococcus* clinical strains from various kinds of purulent wounds (Slopek *et al.*, 1981; Wright *et al.*, 2009; Maura and Debarbieux, 2011). These viruses developed also the ability

to tunnel through bacterial biofilms by employment of specific enzymes, which degrade the bacterial exopolysaccharides (EPS), one of the main component of the biofilm matrix. Such phages can be identified by the appearance of a halo zone around the phage plaques, which results from the enzymatic degradation of bacterial EPS without phage infection (Azeredo *et al.*, 2008). Furthermore, bacteriophages can kill bacteria with peptidoglycan degrading enzymes, endolysins and virion-associated peptidoglycan hydrolases (VAPGH), that disrupt specific bonds in the peptidoglycan, a major component of bacterial cell wall. The use of bacteriophages and their recombinantly manufactured proteins offers a great opportunity to bypass antibiotic therapy hurdles. This dissertation specifically focuses on bacteriophages and their two types of enzymes: the polysaccharide depolymerases and peptidoglycan degrading enzymes, endolysins and VAPGHs, active against *P. aeruginosa.*

In first part of this study we focused on characterization of four newly isolated bacteriophages KT28, KTN6, KTN4 and PA5oct, lytic towards *P. aeruginosa.* Their basic biology was evaluated, including morphology, host surface receptors, host range, stability and infection process. All phages were stable in a range of temperatures, pH and in presence of chloroform. KT28, KTN6 and KTN4 had relatively broad host range, in contrast to PA5oct. Host cell receptors were different for each type of phage, and included LPS, type IV pili and flagella. Furthermore, the phage genomes were sequenced and analyzed in depth. Unwanted genes or mobile elements were not found, which supports obligatory lytic nature of these bacteriophages. Functional genome analysis was supported with ESI-MS/MS analysis of phage proteome and RNA seq. With the use of comparative genomics, the protein-sharing network was constructed, that revealed phage evolution and homology. KTN6 and KT28 belong to *Pbunavirus*, KTN4 is a *Phikzvirus* and PA5oct is a unique giant virus. Finally phage antibacterial activity was evaluated *in vitro* using a novel Airway Surface Liquid model on non-CF and CF epithelial cells lines, in an effort to mimic *in vivo* conditions of the respiratory tract, developed at the laboratory of Prof. B. Harvey (Department of Molecular Medicine, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland). Phages KTN4 and PA5oct presented substantial antibacterial activity, reducing bacterial load from 2.5 to 7 logs, depending on the *P. aeruginosa* strain used.

Second part of this dissertation focuses on phage-encoded PG degrading enzymes (KT28 gp49, KT28 gp41 and its domain, KTN6 gp46, PA5oct gp214, PA5oct gp250, KTN4 gp48) and polysaccharide depolymerases (LKA1 gp49, LUZ7 gp56 and their domains). Their secondary and tertiary structure was analyzed in depth based

on available homology and crystal structures. The corresponding recombinant proteins have been produced and their bactericidal activity and biofilm eradication potential was evaluated. Among the PG degrading enzymes included in this study, the strongest antibacterial activity presented KTN6 gp46, a globular endolysin. Research of polysaccharide depolymerases was complicated, time consuming and required the greatest effort to obtain bioactive, recombinant and purified protein. LKA1 gp49 and its domain reduced EPS slime surrounding bacterial cell. In the future, their biofilm degradation activity should be further evaluated with more specific and quantitative biofilm assays.

This research presents the potential of bacteriophages and their enzymes, which can be considered for therapy or industrial purposes. However, further research is required to analyze e.g. antibacterial activity in details, stability, dose, toxicity or structure. Furthermore, *in vivo* animal studies would show the true power of this alternative therapy.

Streszczenie

Wzrost oporności na antybiotyki pośród bakterii chorobotwórczych stanowi poważny problem terapeutyczny na całym świecie. Coraz częściej pojawiają się wielolekooporne infekcje bakteryjne, które prowadzą do nasilonych i przedłużających się objawów oraz powikłań (Gould, 2006; Haecker, 2009;. Hübner i wsp., 2012). Tlenowa Gram-ujemna bakteria z gatunku *Pseudomonas aeruginosa* jest jednym z najważniejszych i najgroźniejszych drobnoustrojów, zasiedlających środowisko szpitalne. Wykazuje ona zdolność do przystosowania się i zasiedlania różnorodnych nisz ekologicznych, od środowisk wodnych i gleby po tkanki roślinne i zwierzęce. Dodatkowo posiada szeroki wachlarz czynników wirulencji, których efektem działania są nie tylko rozległe uszkodzenia tkanek, ale również zaburzenie funkcjonowania układu odpornościowego (Dzierżanowska, 2008). *P. aeruginosa* jest oportunistycznym patogenem, który powoduje endogenne infekcje u osób z niedoborami odporności (AIDS, mukowiscydoza, nowotwór), a jego eradykacja jest trudna ze względu na dużą oporność na antybiotyki (Cornelis i wsp., 2008).

W trakcie przedłużającej się infekcji populacja *P. aeruginosa* ulega szybkim zmianom zarówno o podłożu genetycznym, jak i fenotypowym, takim jak redukcja produkcji czynników wirulencji oraz zmiana formy życia z planktonicznej na wielokomórkowe populacje, zwane biofilmem. Biofilm odgrywa ważną rolę w ewolucji odporności na antybiotyki i chroni bakterie przed niekorzystnymi warunkami środowiska, w tym mechanizmami obronnymi gospodarza. Hiperzmienność fenotypowa szczepów *Pseudomonas* izolowanych od pacjentów z mukowiscydozą, dodatkowo wpływa na trudności w diagnostyce i leczeniu (Winstanley i wsp., 2016).

W związku z tym zapotrzebowanie na alternatywne metody leczenia jest duże. Pomysł wykorzystania bakteriofagów, naturalnych drapieżników bakterii, jest dobrze znany. Wirusy te wspólnie ewoluowały ze swoim gospodarzem bakteryjnym, w związku z tym wytworzyły bardzo wydajne mechanizmy przeciwbakteryjne, które stanowią ciekawą alternatywę w stosunku do klasycznych antybiotyków. Koktajle fagowe są stosowane jako zastępcza lub wspomagająca leczenie antybiotykowe terapia w krajach Europy Wschodniej. Pozytywne rezultaty uzyskano przy zwalczaniu infekcji ran, wywołanych przez bakterie z rodzaju *Escherichia, Pseudomonas, Proteus, Klebsiella* i *Staphylococcus* (Ślopek i wsp, 1981;. Wright i wsp., 2009, Maura i Debarbieux 2011). Dodatkowo wirusy te posiadają zdolność rozprzestrzeniania się w strukturze biofilmu

poprzez zastosowanie specyficznych enzymów, które degradują egzopolisacharydy bakteryjne (EPS), główny składnik macierzy biofilmu. Tego typu fagi można zidentyfikować poprzez obecność strefy halo wokół łysinki fagowej, utworzonej na murawie bakteryjnej, które to halo jest wynikiem degradacji enzymatycznej bakteryjnego EPS (Azeredo i wsp., 2008). Bakteriofagi wykorzystują jeszcze inne enzymy zwane endolizynami oraz hydrolazami peptydoglikanu związanych z wirionem (VAPGH). Białka te degradują specyficzne wiązania w peptydoglikanie, głównym składniku ściany komórkowej bakterii. Fagi i ich rekombinowane białka dają wiele możliwości zastosowania ich jako preparatów praciwbakteryjnych wspomagających klasyczną antybiotykoterapię. Niniejsza rozprawa doktorska skupia się na charakterystyce bakteriofagów oraz kodowanych przez nie dwóch typach enzymów, depolimerazach polisacharydów oraz endolizynach i VAPGH, aktywnych wobec *P. aeruginosa*.

Pierwsza część prezentowanej pracy badawczej skupia się na charakterystyce czterech nowo izolowanych bakteriofagów KT28, KTN6, KTN4 i PA5oct, litycznych wobec *P. aeruginosa*. Przebadano ich podstawową biologię, ze szczególnym uwzględnieniem morfologii cząsteczki wirusowej, identyfikacji receptorów na powierzchni komórki gospodarza, zakresu litycznego, stabilności wirionów oraz procesu infekcji. Wszystkie fagi wykazały się stabilnością w szerokim zakresie temperatury, pH oraz w obecności chloroformu. Wszystkie, za wyjątkiem PA5oct, miały stosunkowo szeroki zakres lityczny. Pośród receptorów powierzchniowych gospodarza bakteryjnego, zidentyfikowano LPS oraz fimbrie typu IV. W dalszej części genomy czterech bakteriofagów zostały zsekwencjonowane i poddane szczegółowej analizie. Ruchome elementy genetyczne oraz integrazy nie zostały zidentyfikowane, co wskazuje na lityczny charakter badanych wirusów. Analiza funkcjonalna genomu została również wsparta sekwencjonowaniem RNA oraz analizą zestawu białek strukturalnych przy użyciu metody ESI-MS/MS. Dzięki zastosowaniu genomiki porównawczej, została również skonstruowana sieć powiązań, która pozwoliła domniemać proces ewolucji fagów oraz ich homologię. Fagi KTN6 i KT28 należą do rodzaju *Pbunavirus*, KTN4 do *Phikzvirus*, natomiast PA5oct należy do nowej grupy wirusów typu gigant. Oceniono również ich zdolność antybakteryjną *in vitro* przy użyciu nowego modelu infekcji dróg oddechowych z zastosowaniem linii komórkowych nabłonka płuc, izolowanych od osób zdrowych oraz chorych na mukowiscydozę. Model ten został opracowany w laboratorium Prof. B. Harvey (Department of Molecular Medicine, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Irlandia) i imituje środowisko

oraz warunki panujące w drogach oddechowych *in vivo*. Fagi KTN4 i PA5oct wykazały się znaczącą aktywnością przeciwbakteryjną, redukując liczbę komórek bakteryjnych w zakresie 2,5 - 7 rzędów cfu/ml, w zależności od badanego szczepu *P. aeruginosa*.

Druga część prezentowanej rozprawy koncentruje się na enzymach degradujących bakteryjny peptydoglikan (KT28 gp49, KT28 gp41 i jego domenie, KTN6 gp46, PA5oct gp214, PA5oct gp250, KTN4 gp48) oraz depolimerazach polisacharydów (LKA1 gp49 LUZ7 gp56 i ich domenach). Druga i czwartorzędowa struktura białkowa została szczegółowo omówiona w oparciu o dostępną homologię sekwencji aminokwasowej oraz dostępne struktury krystaliczne białka. Czyste białka rekombinacyjne zostały wyprodukowane i oceniono ich potencjał bakteriobójczy oraz zdolność do degradacji biofilmu. Pośród enzymów degradujących peptydoglikan najsilniejsze działanie przeciwbakteryjne wykazało białko KTN6 gp46 o globularnej strukturze. Badania depolimeraz polisacharydów okazały się skomplikowane, czasochłonne i wymagały najwięcej nakładu pracy, w celu uzyskania czystych bioaktywnych białek rekombinacyjnych. Wykazano, że LKA1 gp49 i jego domena były zdolne do redukcji zewnątrzkomórkowego śluzu EPS. Aczkolwiek dalsza analiza jest wymagana, w celu dokładnego zbadania ich aktywność wobec biofilmu, przy zastosowaniu bardziej specyficznych metod.

Przedstawione badania obrazują potencjał bakteriofagów i ich enzymów w leczeniu oraz w przemyśle. Niemniej jednak dalsze badania są wymagane, w celu szczegółowej analizy aktywności przeciwbakteryjnej, stabilności cząsteczek wirusowych i ich białek, potencjalnej toksyczności oraz struktury białkowej. Dodatkowo, badania *in vivo* na modelach zwierzęcych pozwolą zweryfikować prawdziwy potencjał tej alternatywnej terapii.

Samenvatting

De verspreiding van antibioticumresistentie bij bacteriële infecties is een opkomend wereldwijd probleem. Aangezien antibioticumresistente bacteriën niet reageren op de bestaande antimicrobiële therapieën, zijn deze infecties ook zwaarder en veroorzaken ze meer complicaties, hetgeen leidt tot langere ziektebeelden en een hogere behandelingskost (Gould, 2006; Haecker, 2009; Hübner *et al.*, 2012). De aerobe Gram-negatieve bacterie *Pseudomonas aeruginosa* is één van de belangrijkste en meest gevaarlijke pathogenen in de ziekenhuisomgeving. Deze bacterie heft de capaciteit om tal van verschillende niches te koloniseren en past zich aan tal van ecologische niches aan, van water en grond omgevingen tot plant- en dierweefsels. *P. aeruginosa* bezit ook een uitgebreid arsenaal aan virulentiefactoren die niet alleen uitgebreide weefselschade geven, maar ook interfereren met het menselijk immuunsysteem (Dzierżanowska, 2008). Als opportunistisch pathogeen staat deze bacterie ook bekend voor het induceren van endogene infecties bij immune-deficiënte patiënten (AIDS, mucoviscidose, kanker) en is bovendien resistent tegen tal van antibiotica (Cornelis *et al.*, 2008).

'Cystic fibrosis' of mucoviscidose is een veelvoorkomende genetische ziekte die wordt veroorzaakt door mutaties in het '**C**ystic **F**ibrosis **T**ransmembrane **C**onductance **R**egulator' gen. De meeste mucoviscidose patiënten sterven rond de gemiddelde leeftijd van 40 jaar, als gevolg van longfibrose, veroorzaakt door de continue aanvallen op het immuunsysteem, veroorzaakt door de chronische pulmonaire infecties door bacteriën (Carmody *et al.*, 2010). Opportunistische pathogenen zoals *P. aeruginosa*, *Staphylococcus aureus*, *Haemophilus influenze* en het *Burholderia cepacia* complex, kunnen slechts zeer moeilijk uitgeroeid worden in de longomgeving, door hun natuurlijke capaciteit om zich zeer snel aan te passen. (Hogardt *et al.*, 2010). Een belangrijk probleem bij de behandeling van deze infecties is de klonale heterogeniteit van isolaten, alsook hun capaciteit om extracellulaire polysacharide matrices (zoals biofilms) te vormen (Campodonico *et al.*, 2008). Deze biofilmmatrix beschermt de ingebedde bacteriën tegen de externe toediening van antibiotica.

Uit dit alles blijkt dat er een dringende nood is voor nieuwe antibacteriële behandelingsmethodes. Ondanks de moeilijkheden in het antibioticumonderzoek en de ontwikkelingspijplijnen (in het bijzonder bij gramnegatieve bacteriën), onderzoekt de wetenschappelijke gemeenschap alternatieve vormen van antimicrobiële therapieën. Het idee om bacteriofagen, de natuurlijke parasieten van bacteriën, te gebruiken als

hiervoor te gebruiken is bekend. Bacteriofagen hebben gedurende drie miljard jaar co-evolutie ondergaan met hun gastheer. Op die manier hebben er zich zeer efficiënte antibacteriële mechanismen ontwikkeld, die een uniek voordeel kunnen bieden ten opzichte van klassieke antibiotica om bacteriën af te doden. Bacteriofaagcocktails worden in Oost-Europa frequent gebruikt in standaardbereidingen als een alternatief of ondersteunende therapie voor antibiotica. Positieve resultaten worden regelmatig bekomen voor het afdoden van *Escherichia, Pseudomonas, Proteus, Klebsiella* en *Staphylococcus* klinische stammen in verschillende vormen van geïnfecteerde wonden (Slopek *et al.*, 1981a,b; Wright *et al.*, 2009, Maura and Debarbieux, 2011). Deze virussen hebben ook eigenschappen ontwikkeld die hen toelaat bacteriële biofilms te degraderen door middel van specifieke enzymen die de bacteriële exopolysachariden (EPS) afbreken, één van de belangrijkste componenten in de biofilmmatrix. Deze fagen kunnen geïdentificeerd worden door het verschijnen van een halo zone rond de plaques van de faag. Deze halo zone is het resultaat van de enzymatische afbraak van het EPS (Azeredo *et al.*, 2008).

Daarnaast beschikken bacteriofagen ook nog over peptidoglycaan-hydrolyserende enzymen (endolysines en virion-geassocieerde enzymen) die specifieke bindingen in de peptidoglycaan, de belangrijkste component van de bacteriële celwand, hydrolyseren. Het gebruik van bacteriofagen en afgeleide enzymen die recombinant worden aangemaakt bieden een belangrijk potentieel om de bestaande barrières in de ontwikkeling van nieuwe antibacteriële therapieën te overbruggen. In dit eindwerk wordt specifiek de nadruk gelegd op *Pseudomonas aeruginosa* en haar bacteriofagen & afgeleide enzymen: de polysaccharide depolymerasen en de peptidoglycaan-degraderende enzymen.

In het eerste luik van deze studie wordt er gefocust op de karakterisering van vier nieuw-geïsoleerde fagen KT28, KTN6, KTN4 en PA5oct, die alle op een strikt lytische manier inwerken op *P. aeruginosa.* De basis microbiologische eigenschappen worden hierbij bestudeerd, waaronder de morfologie, gastheerreceptoren, gastheerspectrum en hun infectieproces. Alle fagen bleken stabiel in een breed spectrum van temperaturen, pH en bij aanwezigheid van chloroform. KT28, KTN6 en KTN4 hebben een relatief breed gastheerspectrum, in tegenstelling tot faag PA5oct. De gastheerreceptoren, waaronder de LPS, type IV pili en flagellen, verschilden van faag tot faag. Daarnaast werd het genoom van deze fagen bepaald en geanalyseerd. Ongewenste genen of mobiele elementen worden niet gevonden, hetgeen het strikt lytische karakter van deze fagen ondersteund. De experimentele functionele genoomanalyse gebeurde met behulp van een ESI-MS/MS analyse en RNAseq. Aan de hand van vergelijkende genomica en het opbouwen van een

proteïn-familie netwerk, kon de faagevolutie en homologie bestudeerd worden. KTN6 en KT28 zijn leden van het genus *Pbunavirus*, KTN4 is een lid van de *Phikzvirus* terwijl faag PA5oct een uniek nieuw virus is. In een laatste stap werd de antibacteriële activiteit van deze fagen *in vitro* geëvalueerd door middel van een nieuw "Airway Surface Liquid model" op niet-CF en CF epitheliale cellijnen, een model ontwikkeld aan het lab van Prof. B. Harvey, als een poging om de *in vivo* condities van de luchtwegen na te bootsen. Fagen KTN4 en PA5oct vertoonden een aanzienlijke antibacteriële activiteit: een reductie van 2.5 tot 7 log eenheden, afhankelijk van de gebruikte *P. aeruginosa* stam.

Het tweede luik van dit doctoraat focust zich op de faag-gecodeerde PG degraderende enzymes (KT28 gp49, KT28 gp41 en een afgeleid domein, KTN6 gp46, PA5oct gp214, PA5oct gp250, KTN4 gp48) en op de polysacharide depolymerasen (LKA1 gp49, LUZ7 gp56 en hun subdomeinen). De secundaire en tertiaire structuur van deze enzymen werd bestudeerd op basis van homologie en beschikbare kristalstructuren. De overeenkomstige recombinante eiwitten werden aangemaakt en hun antibacteriële en antibiofilm activiteit werd geëvalueerd. Van de geanalyseerde PG hydrolasen vertoonde KTN6 gp46 (een globulair eiwit) de sterkste antibacteriële activiteit. Het onderzoek naar de polysacharide depolymerases bleek ingewikkeld, tijdrovend en vereiste specifieke optimalisaties om recombinant, opgezuiverd en bioactief enzyme te bekomen. LKA1 gp49 en een domein van dit eiwit konden het viskeuze EPS en LPS rond de bacteriële cel. In de toekomst moet deze biofilm degradatieactiviteit verder bestudeerd worden en zijn meer kwantitatieve assays nodig om dit te bereiken.

Dit onderzoek toont het potentieel van bacteriofagen en hun enzymen voor therapeutische en industriële doeleinden. Er is echter nog verder onderzoek nodig, bv. om de antibacteriële activiteit verder te bestuderen, de stabiliteit, dosering, toxiciteit en structuur. Een verdere focus op *in vivo* dierproeven zijn nu aan de orde om de kracht van deze alternatieve therapie te tonen.

List of abbreviations

3D	3-dimensional
A	deoxyadenosine
aa	amino acids
AIDS	Acquired Immune Deficiency Syndrome
AlgT(U)	gene controlling alginate gene expression
Amp	ampicillin
APEC	avian pathogenic *Escherichia coli*
APS	ammonium persulfate
ARG	arginine
ASL	airway surface liquid
ASN	asparagine
ASP	aspartic acid
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
bp	base pairs
CBD	cell wall-binding domain
CCWP	carbohydrate cell wall polymer
OD	optical density
c-di-GMP	cyclic diguanylate
CF	cystic fibrosis
cfu/ml	colony forming units per milliliter
CFTR	transmembrane conductance regulator
Cif	CFTR inhibitory factor
Cm	chloramphenicol
CM	cytoplasmic membrane
DTR	direct terminal repeat
cup	fimbrial chaperone/usher pathway
Cyt	cytoplasm
EAD	enzymatically active domain
ECL	enhanced chemiluminescence

eDNA	extracellular DNA
EDTA	ethylenediaminetetraacetic acid
ENaC	epithelial sodium channel
endoE	endosialidase E
EPS	exopolysaccharide
ESBL	β-lactamase extended substrate spectrum
ESI-MS/MS	electrospray ionization tandem mass spectrometry
ExoA	gene encoding exonucleases
Fe	iron
Feo system	system responsible for transport of ferrous iron into bacteria
FliD	flagella protein
FPLC	fast protein liquid chromatography
Fur	ferric uptake regulator
FyuA	yersiniabactin receptor gene
GFP	green fluorescent protein
GlcNAc	N-acetylglucosamine
GLN	glutamine
GLY	glycine
Gm	gentamicin
gp	gene product
Has system	hemophore-dependent heme acquisition system
HEWL	hen egg white lysozyme
HIS	histidine
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
ICTV	International Committee on Taxonomy of Viruses
ICUs	intensive care units
IFNγ	gamma interferon
IL-8	interleukin 8
IPTG	isopropyl β-D-1-thiogalactopyranoside
ILE	isoleucine
kbp	kilobase pairs

(k)Da	(kilo)Dalton
λ-DE3	lysogen construct with T7 RNA polymerase under control of lacUV5 promoter
L	liter
LadS/GacS	systems involved in regulation of secondary metabolism and ther aspects of bacterial physiology
LasA	elastase
LasB	elastase
LB	lysogeny broth
LEU	leucine
LMW	low molecular weight
LPS	lipopolysaccharides
LU	linkage unit
M	mol
MBL	metallo-β-lactamase
MCL	Markov clustering algorithm
mDAP	meso-diaminopimelic acid
MET	methionine
MHB	Mueller-Hinton broth
MIC	minimum inhibitory concentration
ml	milliliter
mM	milimol
MOI	multiplicity of infection
MQ	Milli-Q water
MRSA	methicillin-resistant *Staphylococcus aureus*
mucA	anti-sigma factor gene
MurNAc	N-acetylmuramic acid
nm	nanometre
NTA	nitrilotriacetic acid
OM	outer membrane
OMP	outer membrane permeabilizing peptide

OprD	protein family containing bacterial outer membrane porins involved in transport of cationic amino acids, peptides, antibiotics and other compounds
ORF(s)	open reading frame(s)
P	phosphate group
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PCs	protein clusters
PEG	polyethylene glycol
Pfu	plaque forming units
pfu/ml	plaque forming units per ml
PG	peptidoglycan
PHE	phenylalanine
Phu system	*Pseudomonas* haem uptake system
Plc	phospholipase
PQS	*Pseudomonas* quinolone signal
PRO	proline
Psl	polysaccharide synthesis locus
RBS	ribosome binding site
RetS	hybrid sensor kinase-response regulator protein
RNAP	RNA polymerase
ROS	reactive oxygen species
QS	quorum sensing
SAR	signal-arrestrelease
SEM	scanning electron microscope
SCV	small-colony variants
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGNH	serine, glycine, asparagine, histidine
SLT	soluble lytic transglycosylase
SOC	super optimal broth with catabolite repression
SP	signal peptides
T3SS	type III secretion system

TEMED	N,N,N’,N’- tetramethylethylenediamine
TLR	Toll-like receptor
TonB	bacterial outer membrane proteins that bind and transport ferric chelates
TYR	tyrosine
UV	ultraviolet
VAPGH	virion-associated peptidoglycan hydrolase
VC	viral cluster
VNPs	viral nanoparticles

List of publications

Danis-Wlodarczyk Katarzyna, Vandenheuvel Dieter, Jang Ho Bin, Briers Yves, Olszak Tomasz, Arabski Michal, Wasik Slawomir, Drabik Marcin, Higgins Gerard, Tyrrell Jean, Harvey J Brian, Noben Jean-Paul, Lavigne Rob, Drulis-Kawa Zuzanna (2016). A proposed integrated approach for the preclinical evaluation of phage therapy in *Pseudomonas* infections. Scientific Reports; 6:28115. doi: 10.1038/srep28115.

Danis-Wlodarczyk Katarzyna, Olszak Tomasz, Arabski Michal, Wasik Slawomir, Majkowska-Skrobek Grazyna, Augustyniak Daria, Gula Grzegorz, Briers Yves, Jang Ho Bin, Vandenheuvel Dieter, Duda Katarzyna Anna, Lavigne Rob, Drulis-Kawa Zuzanna (2015). Characterization of the newly isolated lytic bacteriophages KTN6 and KT28 and their efficacy against *Pseudomonas aeruginosa* biofilm. PLoS One, 10 (5), doi: 10.1371/journal.pone.0127603, e0127603

Olszak Tomaszs, Paulina Zarnowiec, Kaca Wieslaw, **Danis-Wlodarczyk Katarzyna**, Augustyniak Daria, Drevinek Pavel, de Soyza Anthony, McClean Siobhan, Drulis-Kawa Zuzanna (2015). *In vitro* and *in vivo* antibacterial activity of environmental bacteriophages against *Pseudomonas aeruginosa* strains from cystic fibrosis patients. Applied Microbiology and Biotechnology vol. 99 issue: 14 pages: 6021-33, doi: 10.1007/s00253-015-6492-6

Drulis-Kawa Zuzanna, Olszak Tomasz, **Danis Katarzyna**, Majkowska-Skrobek Grazyna, Ackermann Hans-W (2014). A giant *Pseudomonas* phage from Poland. Archives of Virology vol:159 issue:3 pages:567-72, doi: 10.1007/s00705-013-1844-y

Chapter 1
Introduction and background

1.1. *Pseudomonas aeruginosa*, an opportunistic pathogen

1.1.1. Taxonomy and general characteristics

In 1872 two German scientists, Joseph Schröter (1872) and Ferdinand Cohn (1872), isolated a blue pigmented microorganism, calling it "*Bacterium aeruginosum*" or "*Bacillus aeruginosus*". In 1900, Polish botanist Emil Friedrich August Walter Migula for the first time coined the name *Pseudomonas aeruginosa* in his "System der Bakterien" (1900).

The word *Pseudomonas* means "false unit", from Greek *pseudo*, ψευδο (false) and Latin *monas*, Greek: μονος (a single unit). *Aeruginosa* is from Latin meaning verdigris, "copper rust" or from the Greek prefix *ae*-meaning "old or aged", and the suffix *ruginosa* means wrinkled or bumpy.

The current classification of *P. aeruginosa* is presented in Table 1.1.

Table 1.1: *P. aeruginosa* taxonomy according Uniprot Database (2015)**.**

Kingdom	*Bacteria*
Phylum	*Proteobacteria*
Class	*γ-Proteobacteria*
Order	*Pseudomonadales*
Family	*Pseudomonadaceae*
Genus	*Pseudomonas*
Species	*Pseudomonas aeruginosa*

P. aeruginosa is a motile Gram-negative rod-shaped bacterium, that can be found in a wide range of environmental niches, such as terrestrial and marine environments, as well as associated with plants and animals, including humans (Fig.1.1). This opportunistic pathogen causes chronic and acute infections of burn wounds, as well as respiratory and urinary tracts (Kobayashi *et al.*, 2009). Particularly vulnerable are patients in intensive care units (ICUs) and the immunocompromised (e.g. cancer, AIDS, cystic fibrosis) (Murray *et al.*, 2007; Kobayashi *et al.*, 2009; Trautmann *et al.*, 2005). *P. aeruginosa* possesses a wide array of virulence factors that do not only cause extensive tissue damage, but also directly interfere with the human immune system (Dzierżanowska, 2008).

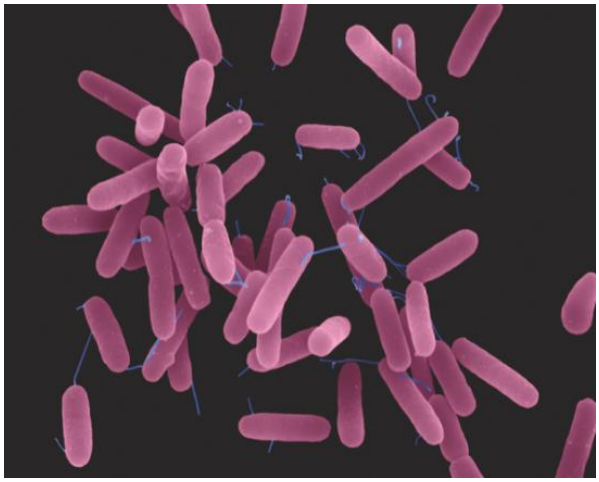


Figure 1.1: *P. aeruginosa* SEM image (© Dennis Kunkel/ Microbiology, Inc.). A rod-shaped bacterium with flagella, dimensions of 0.5 to 0.8 µm by 1.5 to 3.0 µm (Lederberg *et al.*, 2000).

A major advantage to this pathogen is its diverse metabolic competence, that enables it to employ small molecules or complex organic compounds as a carbon source and aerobic as well as anaerobic growth in the presence of NO_3^- as electron acceptor. *P. aeruginosa* has broad growth temperature range (6 - 45°C, with an optimum growth at 37°C) (Slonczewski *et al.*, 2011; Szewczyk, 2005). Furthermore, it lives as planktonic cells, colonies or biofilms, reversibly regulates gene expression and changes its phenotype and genotype in response to the environmental signals during infection (Essoh *et al.*, 2013; Bragonzi *et al.*, 2009). This provides a protective mode that allows survival in hostile environments and disperses seeding cells to colonize new niches with desirable conditions (Wei *et al.*, 2013).

P. aeruginosa cultures can create three different morphological types of colonies: 1) colonies with a diameter ~ 2 mm, rough, gray color, the slightly dull surface with irregular edges (the environmental isolates from soil/water); 2) large colonies, smooth, shiny and flattened at the edges (mostly clinical isolates); 3) mucous with alginate overproduction (clinical isolates, mainly from respiratory and urinary tracks) (Szewczyk, 2005).

Another diagnostic significant characteristic is the *P. aeruginosa* pigment production capacity. These pigments can diffuse through the medium, giving rise to colonies with variable colors (Fig. 1.2). Depending on the strain, these pigments include pyocyanin (blue-green, nonfluorescent), pyorubin (red-brown), pyomelanin (brown-black, rather uncommon) and pyoverdine (yellow-green, fluorescent under UV light) (Khalifa *et al.*, 2011; Jayaseelan *et al.*, 2014).

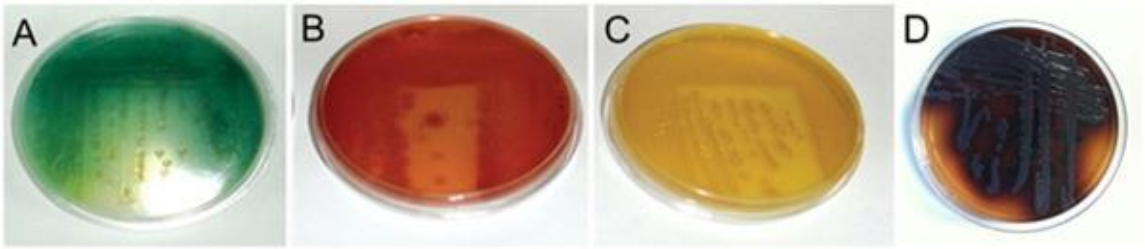


Figure 1.2: Different pigmentation of *P. aeruginosa*. A) *P. aeruginosa* E1 isolated from refinery soil, pyocyanin overproduction B) *P. aeruginosa* C6 isolated from burn wound, pyorubrin overproduction. C) *P. aeruginosa* C10 isolated from burn wound, pyoverdine overproduction. D) *P. aeruginosa* PA1111, clinical isolate, pyomelanine overproduction. The cetrimide medium (Yu *et al.*, 2002; Nowroozi *et al.*, 2012; Ketelboeter *et al.*, 2014).

1.1.2. Pathogenesis of *Pseudomonas aeruginosa*

1.1.2.1. *Pseudomonas aeruginosa* infections, a public health crisis brews

P. aeruginosa belongs to infamous ESKAPE group of pathogens (***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumoniae*, ***A****cinetobacter baumannii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* species), that effectively "escape" the conventional antimicrobial therapies and is considered to be one of the major threats to public health (Rice, 2010; Boucher *et al.*, 2009). This opportunistic pathogen is a leading Gram-negative organism associated with nosocomial infections (11 - 13.8% of total), that can spread e.g. during intubation, tracheotomy, colostomy, operations, cross-contaminations between patients and water fittings (Dzierżanowska, 2008; Driscoll *et al.*, 2007). Infections are associated with a disrupted epithelial layer (e.g. burn wounds) or mucosal barrier (e.g. lung infections), malfunctioning of the immune system (e.g. AIDS) or when bacteria are introduced into the eye (e.g. by contact lenses). *P. aeruginosa* can also form biofilms on the medical devices surface, such as catheters, intubation tubes or implants (Ballok *et al.*, 2013). The mortality range is 15 - 78% in ICUs. Among hospital-acquired *P. aeruginosa* infections, we may distinguish: 1) ventilator-associated pneumonia with ~20% prevalence and dramatically high mortality (Fujitani *et al.*, 2011, Kollef *et al.*, 2014); 2) chronic obstructive pulmonary disease; 3) pneumonia in cystic fibrosis (CF) patients; 4) burn and postoperative/surgical wound infections; 5) urinary tract infections (Table 1.2). During their lifetime, all CF patients experience *P. aeruginosa* lung infections and more than 80% of them become chronically infected in their adult years (Döring *et al.*, 2011). Such a lifelong infection decreases lung function and ultimately results in death (Minandri *et al.*, 2016).

Table 1.2: Most frequent diseases correlated with *P. aeruginosa* infections (Rehm, 2009).

Organ	Infection	Symptoms	Acute/ Chronic	Origin	Prevention
eye	contact lens, keratitis	inflammation pain swelling redness impaired vision	acute	water	hygiene
ear	external otitis	swelling ear pain itching inside the ear discharge from the ear difficulty hearing	acute/ chronic	possibly water	hygiene
nasal sinuses	sinusitis	swelling pain difficulty breathing	chronic	possibly water	unknown
skin	folliculitis, wound/ ulcer infection, burn infections	redness of the skin abscess formation in the skin draining wounds	acute	water, soil	hygiene
urine bladder	urinary tract infection	persistent urge to urinate burning sensation when urinating cloudy and smelling urine	acute/ chronic	nosocomial	early aggressive antibiotic eradication therapy
bones	diabetic osteomyelitis in feet	swelling pain in the affected joint	chronic	possibly water	hygiene
blood	sepsis, neutropenic patients	fever and chills fatigue muscle and joint pain hemodynamic shock	acute	unknown	antibiotic prophylaxis
lungs/ bronchi	ventilator-associated pneumonia	chills fever cough with or without sputum production difficulty breathing	acute	nosocomial humidifiers	hygiene
	endobroncholitis, cystic fibrosis, bronchiectasis		chronic	nosocomial, environment	hygiene, early aggressive antibiotic therapy of intermittent colonization

Infection in CF environment

The best known genetically inherited diseases with increased prevalence of *P. aeruginosa* infections is CF, correlated with defects in a transport protein, the CF transmembrane conductance regulator (CFTR) (Winstanley *et al.*, 2016). CFTR is a chloride ion channel of the ABC transporter family and its mutation result in misfolding, a lack of proper localization, and/or a complete lack of the protein (Chambers *et al.*, 2007). In cooperation with the epithelial sodium channel (ENaC), this protein is responsible for controlling the level of airway surface liquid (ASL) layer, that covers the lung surface and consists of a mucus and a periciliary liquid layer. The former traps and

removes inhaled pathogens, while the latter keeps the mucus at an optimal distance from the underlying epithelia to maximize ciliary mobility, provides a low viscosity solution and acts as a lubricant layer for mucus transport (Danis-Wlodarczyk *et al.*, 2015). Together, they play a critical role in effective mucociliary clearance of the airway (Matsui *et al*., 1998;. Zabner *et al*., 2003). In the absence of CFTR, ENaC is hyperactive leading to loss of a chloride secretion and increase in a sodium import. This results in the decreased ASL volume or altered mucus impede phagocyte function, making bacterial elimination defective. Moreover, inflammatory response is abnormal (Danis-Wlodarczyk *et al.*, 2015). The CF airways sticky mucus provides the perfect milieu, microaerophilic to anaerobic environment, enabling the colonization and propagation of *P. aeruginosa* (Worlitzsch *et al.*, 2002; Danis-Wlodarczyk *et al.*, 2015). The resident microbial community of CF lungs is known to be complex and it has considerably changed, mainly due to alterations in antibiotic regimens. Nevertheless, *P. aeruginosa* is still the most prevalent pathogen isolated from CF sputum, occurring more frequently in adults (Folkesson *et al.*, 2012; Sousa *et al.*, 2014). Once entering in CF airways, *P. aeruginosa* is practically impossible to eradicate due to its remarkable genome plasticity, that allows it to rapidly adapt to the specific CF environment (Hogardt *et al.*, 2010).

Establishing an infection

Most CF patients acquired pathogens mainly from environment, especially in clinical settings where CF patients remain for long periods of time. *P. aeruginosa* has to properly regulate its gene expression to quickly adapt to a hostile CF lung environment, where it is exposed to host immune defenses, antibiotics, and oxidative stress. The bacterial features among CF acute isolates fluctuate significantly. However, there is a trend towards high virulence potential and cytotoxicity as well as lower frequency of mutator strains. The expression of virulence factors, including cell-associated and secreted virulence factors, is considered to be fundamental at early stage for the success of infection establishment (Sousa *et al.*, 2014). A first group of virulence factors includes the flagellum, type IV pili, lipopolysaccharide (LPS) and alginate. A second group includes e.g. elastases (LasB and LasA), alkaline protease, pyocyanin, and the effector molecules of the type III secretion system (T3SS) (Qaisar *et al.*, 2016). Flagella, pili, and LPS are important for motility and adhesion by binding to respiratory mucin and the glycolipid, asialoGM1, as well as for activations of Toll-like receptors (TLR5, TLR2, TLR4), which leads to immune activation (McIsaac *et al.*, 2012; Amiel *et al.*, 2010; Lau *et al.*, 2005).

The secretion of protein ExoA gives *P. aeruginosa* the ability to stop epithelial cell protein expression and kill host cells (Wolf *et al.*, 2009; Deng *et al.*, 2008; Liu, 1973). The T3SS plays a major role in establishing infection (Bleves *et al.*, 2010; Engel *et al.*, 2009). The T3SS effectors ExoS, ExoT, ExoY, ExoU inhibit actin polymerization, prevent phagocytosis and cell migration, and promote apoptosis (Barbieri *et al.*, 2004). They also modulate the inflammatory response, increase membrane permeability (Ochoa *et al.*, 2012), leading to membrane damage, and cell lysis (Deng *et al.*, 2008; Anderson *et al.*, 2012).

At an early stage of infection, eradication is still possible when a suitable antibiotic treatment has been started in a timely manner. Otherwise, 20% of those first *P. aeruginosa* colonization could become directly chronic infections and may persist up to the end of the patient's life (Bragonzi *et al.*, 2009; Montanari *et al.*, 2007; Sousa *et al.*, 2014).

Internalization in lung epithelia cells

Although *P. aeruginosa* was generally thought to be an extracellular pathogen, a number of different groups have found that it can be internalized into a range of different cell types, including epithelial cells (Darling *et al.* 2004; Fleiszig *et al.*, 1996 and 1995; Evans *et al.*, 1998; Plotkowski *et al*., 1999; Burns *et al.*, 1996; Danis-Wlodarczyk *et al.*, 2015). Fleiszig *et al.* (1996) results suggested, that the determination as to whether a strain of *P. aeruginosa* is invasive in epithelial cells is related to its ability to kill eukaryotic cells. Strains with low cytotoxicity showing high levels of invasion and vice versa. Schroeder *et al.* (2001) presented that the highest attachment of PAO1 was after 3 h of infection and the maximum percentage of internalized bacteria was found after 6 h of incubation. However, it is not clear how the different mechanisms resulting in *Pseudomonas* internalization are switched on, how such a decision is established at the level of bacteria and host cells, and which structures govern this process (Byfield *et al.*, 2011).

To start the invasion of the lung epithelial cells, *Pseudomonas* first has to find its receptor and adhere to the cell surface. In healthy epithelum the CFTR protein serves as a cell surface receptor for non-mucoid smooth strains of *P. aeruginosa* with an intact LPS core (Pier *et al.*, 1997; 1996a, b). However, Darling *et al.* (2004) stated that CFTR does not exert a direct stoichiometric effect on *P. aeruginosa* uptake in a specific cell, but rather alters the properties of the whole apical epithelial surface to inhibit bacterial invasion. The CFTR does not occur with mucoid rough *P. aeruginosa* strains that emerge during lung colonization in CF patients (Pier *et al.*, 1996b). A study by Kato *et al.* (2010) showed that also MUC1 mucin serves as a binding site for *P. aeruginosa*. The adhesion

process is supported by type IV pili, Psl exopolysaccharide (EPS) and flagella FliD protein that binds to mucins (Arora *et al.*,1998; Häussler, 2003; Merritt *et al.*, 2007). In CF *P. aeruginosa* internalize without epithelial cell apoptosis, resulting in intracellular bacterial survival, division, and escape from the host immune system (Grassme *et al.*, 2000; Byfield *et al.* ,2011).

Infection maintenance

Once *P. aeruginosa* has succeeded in lung invasion and inhibited clearance, it must induce changes in order to reduce immune activation and obstruct clearance mechanisms to persist in the lung (Ballok *et al.*, 2013). To facilitate mucus penetration, the pathogen employs a suite of secreted enzymes (exoproteins) to impair host immune system (Bleves *et al.*, 2010; Lau *et al.*, 2005; Ballok *et al.*, 2013). These immunosuppressing factors comprise the elastases LasA and LasB, and phospholipases PlcB, PlcH and PlcN, which target the mucus layer and cell membrane, facilitating bacterial transit through the mucus layer and liberating nutrients exploited by the bacteria (Ochsner *et al.*, 2002; Schmiel *et al.*, 1999; Barker *et al.*, 2004). Furthermore, the neutrophil respiratory burst is suppress by PlcH (Terada *et al.*, 1999). The CFTR inhibitory factor (Cif), an epoxide hydrolase, reduce apical membrane CFTR and Cl^- secretion by epithelial cells as well as a drug efflux pump, P-glycoprotein, inhibiting the microbicidal activity of neutrophils by limiting Cl^- import to endosomes and preventing hypochlorous acid formation in *P. aeruginosa*-containing vesicles (Painter *et al.*, 2008; Bonvillain *et al.*, 2010). A high expression of the secreted alkaline protease (AprA) correlates with mucoidy aid in *P. aeruginosa* survival by cleaving transferrin to facilitate iron acquisition by siderophores (Firoved *et al.*, 2003; Jaffar-Bandjee *et al.*, 1995; Kim *et al.*, 2006), as well as inhibiting immune recognition by TLR5 and degradation of complement proteins and gamma interferon (IFNγ), halting phagocytosis and killing by neutrophils (Laarman *et al.*, 2012; Butterworth *et al.*, 2012; Ballok *et al.*, 2013). Along with these extracellular proteins, in many CF strains production of the EPS alginate increases, which promotes mucin production, thereby limiting immune recognition and clearance (Kishioka *et al.*, 1999; Ohman *et al.*, 1981; Deretic *et al.*, 1987).

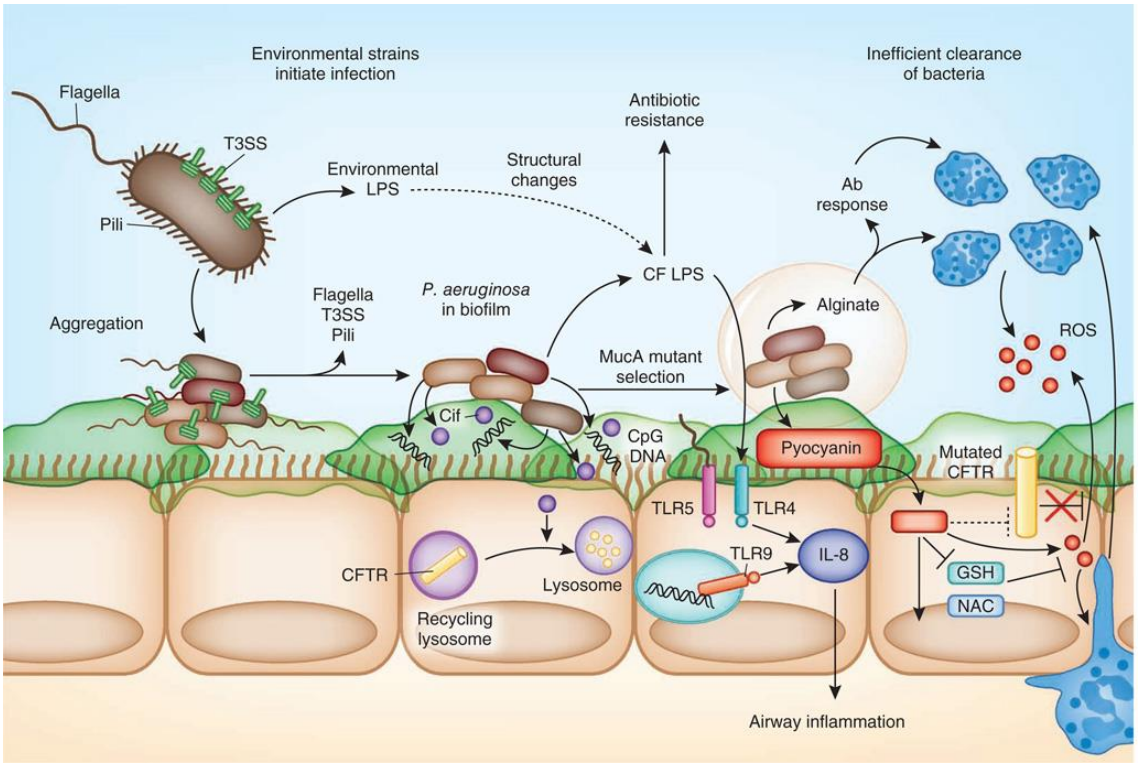


Figure 1.3: Development of chronic infection of *P. aeruginosa* in the cystic fibrosis airway. Inhaled pathogens expressing flagella, pili and T3SS aggregate within the lung, resulting in the formation of biofilm. Within the biofilm, bacteria lose flagella, pili and the T3SS, increase alginate production and express a diverse range of virulence factors promoting evasion of the host immune system. *P. aeruginosa* also releases Cif in outer membrane vesicles to inhibit the recycling of CFTR in the host. Furthermore, the lipid A structure of the LPS is altered through the addition of palmitate and aminoarabinose, resulting in increased antibiotic resistance and increased induction of IL-8 production by host cells (Cohen *et al.*, 2012).

Phenotypes and genotypes selection during persistence in the CF lung

The adaptation process to CF airways can be accelerated by the emergence of mutator phenotypes (or hypermutable phenotypes), that can have mutation rates up to 1000–fold in compare to non-mutator phenotypes (Oliver *et al.*, 2000; Rodriguez-Rojas *et al.*, 2012; Mena *et al.*, 2008). In the hostile, extremely selective conditions of the CF airways, this sophisticated mechanism improves *P. aeruginosa* microevolution and its intraclonal diversification. The development of phenotypic variants and mutators can be intrinsic, relying on mutations or recombinations generated by defects on DNA repair or error avoidance systems (Sousa *et al.*, 2014). Mutators can also be influenced by environmental factors, e.g. reactive oxygen species (ROS) generated from inflammatory responses, that can trigger the generation of phenotypic variants damaging DNA and cause mutations in bacteria (Ciofu *et al.*, 2005; Sousa *et al.*, 2014). Furthermore, the antibiotic sub-inhibitory or sub-lethal concentrations can also induce mutations and recombinations (Kohanski *et al.*, 2010; Sousa *et al.*, 2014).

Despite the selective advantage of mucoid *P. aeruginosa*, isolates from a single patient may differ greatly in their morphotypes including mucoid, smooth, rough, dwarf, non-pigmented variants, small-colony variants, and variants with visible autolysis or autoaggregative behavior that further may all differ in their antibiotic susceptibility patterns (clonal diversification to different niches in CF airways) (Hogardt *et al.*, 2007; Sousa *et al.*, 2014). Non-mucoid isolates may be either wild-type *P. aeruginosa* or revertants from mucoid phenotypes. The mucoid phenotype is unstable during ongoing CF lung disease, probable due to high energy requirement for alginate production (Sousa *et al.*, 2014).

Mucoid *P. aeruginosa* isolates typically lack twitching and swimming motility due to non-piliation and loss of the flagellum, probably since motility is not needed during static, non-gradient growth conditions within anaerobic biofilms and might enhance survival by decreasing neutrophil recruitment and phagocytosis by alveolar macrophages and granulocytes in the airway (Lau *et al.*, 2005; Hajjar *et al.*, 2002). Moreover, chronic *P. aeruginosa* CF isolates typically show a rough colony phenotype due to modification of the lipid A moiety of LPS (Hancock *et al.*, 1983), associated with increased resistance to antimicrobial peptides. Further, the EPS alginate is overproduce, which is widely considered to be a marker for the transition to chronic infection (Winstanley *et al.*, 2016).

Small-colony variants (SCV) are often selected for by prolonged antibiotic exposure, common during chronic infections. They have a much slower growth rate, smaller colony morphology and exhibit normally hyperpiliated, hyperadherent and autoaggregative behavior which can favor biofilm formation and persistence (Cullen *et al*., 2015). Further, it has been demonstrated that the increased expression of the pel and psl EPS gene clusters and elevated intracellular cyclic di-GMP (c-di-GMP) levels contribute to SCV morphotype (Starkey *et al.*, 2009).

1.1.2.2. Quorum sensing

The quorum sensing (QS) is a coordination mechanism, responsible for regulation of gene expression of many virulence factors (Miller *et al.*, 2001; Winstanley *et al.*, 2009). Bacteria produce and release autoinducers, a small diffusing molecules, which concentration increases as function of cell density. When minimal stimulatory concentration threshold of these autoinducers is detected, an alteration in gene expression occur (Miller *et al.*, 2001). The quorum sensing is a form of communication between microorganisms, which optimizes their metabolism and behavior while living in dense

colonies. The switch from planktonic to sessile lifestyle depends on two conflicting sensor systems: 1) RetS, responsible for initiation of acute virulence with the production of toxins and expression of T3SS; 2) LadS/GacS, that leads to chronic infections by selection of the sessile mode with the expression of genes involved in type VI secretion system, EPS production and biofilm formation (Coggan *et al.*, 2012; Balasubramanian *et al.*, 2013; Cornelis *et al.*, 2013; Jimenez *et al.*, 2012; Sousa *et al.*, 2014).

1.1.2.3. Iron deprivation

In the human body, iron ions are not readily available due to iron sequestration by different iron-binding proteins, including transferrin, lactoferrin and ferritin (Wang *et al.*, 2011; Qaisar *et al.*, 2016; Minandri *et al.*, 2016). To survive iron deprivation *P. aeruginosa* utilizes multiple systems and strategies to obtain iron from its environment: 1) siderophores, an extracellular Fe^{3+} chelating molecules, including pyoverdine and pyochelin with high and low affinities for iron respectively, and the uptake of ferrisiderophores via TonB-dependent receptors; 2) the uptake of xenosiderophores-siderophore piracy from other microorganisms (e.g. *E. coli* enterobactin); 3) the heme molecule uptake from the host hemoproteins, through the Has and Phu systems; 4) the extracellular reduction of Fe^{3+} to Fe^{2+} involving phenazine compounds and a Fe^{2+} dedicated iron uptake system under microaerobic or anaerobic conditions (the TonB-independent Feo system) (Hamood *et al.*, 2004; Anzaldi *et al.*, 2010; Cornelis, 2013; Nairz *et al.*, 2010; Cox, 1993; Minandri *et al.*, 2016). The iron acquisition process is highly regulated and involves multiple regulators including the ferric uptake regulator (Fur) (Cornelis *et al.*, 2009), a repressor of iron uptake genes when bound to its co-repressor Fe^{2+} (Escolar *et al.*, 1999). Qaisar *et al.*(2016) showed, that the *pvcA-D* operon, indirectly through a new discovered virulence factor isonitrile functionalized cumarin, paerucumarin, influences biofilm development by enhancing the expression of fimbrial chaperone/usher (*cup*) pathway genes. Paerucumarin as well as *Pseudomonas* quinolone signal (PQS) also chelates iron within the extracellular environment to facilitate the induction of *P. aeruginosa* siderophores, but does not function as a siderophore. When iron is available extracellularly, the level of intracellular iron increases and Fur is activated. Activated Fur then represses the expression of the siderophores. Depending on the type of infection, *P. aeruginosa* can therefore adapt by switching from one iron uptake system to another (Cornelis, 2013).

1.1.2.4. *P. aeruginosa* biofilm and its influence on multidrug resistance

General characteristics

Most bacteria have a propensity to adhere to surfaces to create a structured multicellular communities embedded in a self-produced polymer matrix, that provide a protective mode to hostile or rapidly changing environments (Wei *et al.*, 2013; Battin *et al.*, 2007; Abedon, 2011; Høiby *et al.*, 2010)**.** They occur in natural, industrial and medical environments (Wei *et al.*, 2013; Hughes *et al.*, 1998; Sutherland *et al.*, 2001). In humans biofilms cause chronic infections, due to increased tolerance to antibiotics and disinfectant chemicals, as well as resisting phagocytosis and other components of the immune defense system (Høiby *et al.*, 2010). The metabolism, physiology and phenotype of biofilm-related cells differ from their free-living planktonic counterparts (Wei *et al.*, 2013; Flemming *et al.*, 2010). Moreover, gradients of access to nutrients and oxygen exist through whole biofilms structure. These gradients are associated with decreased bacterial metabolic activity, dormancy and anaerobic conditions in the center of biofilm (Khalifa *et al.*, 2011; Hanlon, 2007; Høiby N *et al.*, 2010). Within a biofilm, *Pseudomonas* cells are in close proximity and strongly interact with each other, including competition (e.g. for nutrients), commensalism (e.g. a metabolic intermediate or waste product of one species can serve as the substrate for another species), mutualism and parasitism. Importantly, the majority of natural biofilms subsist multiple bacterial species and each may implement different specialized tasks (Khalifa *et al.*, 2011).

Biofilms in nature have many different forms, ranging from smooth, flat biofilms to three-dimensional mushroom-shaped structures. The amount of water channels and size of pores between the microcolonies can also vary substantially. In this manner, the mature biofilm resembles primitive, multicellular organisms (Høiby *et al.*, 2010). The shape of the formed biofilm is determined by the motility, which is influenced by the carbon source and QS (Flemming *et al.*, 2010).

Matrix composition

In a biofilm, bacteria are embedded in the extracellular polymeric substances matrix, which contains EPS, extracellular DNA (eDNA), proteins, lipids and ions, which strongly interact with each other and associate in a complex network that holds mainly by electrostatic forces and hydrogen bonds (Flemming *et al.*, 2010; Kaur *et al.*, 2015). The composition and structure depends on many factors, including the component

microbial cells, their physiological state, the physical environment, the available nutrients and substrates, and the surface to which the cells are attached (Kaur *et al.*, 2015). In addition, the chemical environment plays a role in the biofilm structure, e.g. calcium ions can form a bridge between the negatively charged alginate polymers and increase the thickness of biofilms (Parsek *et al.*, 2008).

The matrix is important since it provides protection to the biofilm. It is necessary for adhesion to surfaces, it holds the biofilm cells together and provides mechanical stability**.** Furthermore, it protects the resident cells against environmental stresses (e.g. dehydratation, metals, UV radiation, pH extremities, antimicrobial agents, immunological defense mechanisms, etc.) and acts as a reservoir by providing retention of water, nutrients and other compounds (Wei *et al.*, 2013).

P. aeruginosa exopolysaccharides

At least three exopolysaccharides have been shown to be produced by *P. aeruginosa*, alginate, Psl, and Pel (Colvin *et al*., 2011). In case of non-mucoid strains, Psl and Pel EPS plays a major role in biofilm formation. While in mucoid counterparts alginate is overproduced. Furthermore, alginate-containing biofilms are structurally heterogeneous, whereas biofilms produced by non-mucoid strains are flat and more homogeneous (Flemming *et al.*, 2010). Not all of three polymers are required for biofilm formation. Mutants that lack Psl, alginate or Pel can still form biofilms, whereas mutants that lack both Psl and Pel are not able form biofilms (Ghafoor *et al.*, 2011).

Psl

Polysaccharide synthesis locus (*psl*) enncodes a polysaccharide composed of L-rhamnose, D-mannose and D-galactose monomers, however the exact structure is not yet revealed (Wei *et al.*, 2013). Psl expression is controlled by a two-component regulatory system and autoregulation by stimulation of the second messenger c-di-GMP production (Harmsen *et al.*, 2010). The Psl polysaccharide forms a helical structure around the bacterial cells and is important for adhesion. It increases cell-to-cell interactions which are essential for biofilm formation assembly at the air-liquid interface and maintenance of non-mucoid *P. aeruginosa* strains (Coulon *et al.*, 2010; Harmsen *et al.*, 2010). Psl is responsible for the formation of a fabric-like matrix and it holds the cells closely together, thereby impeding antibiotic penetration throughout the biofilm**.** This EPS is mainly situated in the outer layers of three dimension-structured microcolonies, whereas it is uniformly

distributed in flat biofilms (Ma *et al.*, 2009). Psl can reduce neutrophil phagocytosis and the oxidative response via limiting complement-mediated opsonization (Mishra *et al.*, 2012). It provides a first line of defense against cationic and anionic antibiotic during initial stages of biofilm formation.

Pel

Pel is a positively charged exopolysaccharide composed of partially acetylated 1→4 glycosidic linkages of N-acetylgalactosamine and N-acetylglucosamine, synthesized by a highly conserved, seven-gene operon (*pelA-F*) (Jennings *et al.*, 2015; Wei *et al.*, 2013). The *pel* locus is regulated by a two-component regulatory system and involves c-di-GMP (Harmsen *et al.*, 2010). Pel plays a role in the later stages of biofilm development and is required for solid surface-associated biofilms**.** It imparts a wrinkled colony phenotype, can compensate as an attachment factor in the absence of other adhesins (e.g. type IV pili) and is necessary for the maintenance of cell-cell interactions in mature biofilms (Friedman *et al.*, 2004). Coulon *et al.* (2010) studies suggesting that *pel* locus might be also involved in the LPS production.

Alginate

Alginate is a linear co-polymeric EPS composed of non-repetitive monomers of β-1,4-linked L-guluronic and D-mannuronic acids and is associated with mucoidy of *P. aeruginosa* (Lau *et al.*, 2005). It is a polyanionic polymer, that provides a physical barrier for some antibiotics (e.g. aminoglycosides) by binding them to its positively charged molecules (Yan *et al.*, 2014; Khalifa *et al.*, 2011). Most mucoid isolates have mutations in the anti-sigma factor gene, mucA, that leads to activation genes involved in alginate biosynthesis by AlgT(U). These polysaccharides have a large molecular mass and are relatively soluble, resulting in the formation of highly viscous aqueous solutions (Sutherland, 2001). Furthermore, alginate plays important role in structural stability of biofilms and protects against human antibacterial defense mechanisms such as phagocytosis, antibodies and ciliary movements (Ma *et al.*, 2009).

Other biofilm components

Up to 97% of the biofilm matrix is composed from water (Wei *et al.*, 2013), that can exist as a solvent for solutes or bound within the capsules of microbial cells. Three-dimensional biofilms may contain water channels, that allow transfer of oxygen, nutrients,

substrate, metabolites, waste products and other solutes throughout the biofilm (Sutherland, 2001).

The matrix also contains nucleic acids, mainly in the form of eDNA, found primarily within the stalks of mushroom-shaped microcolonies. eDNAs are random chromosomal DNA segments that are either actively secreted in small membrane vesicles or released from cells after lysis caused by bacteriocins, bacteriophages or QS. For migrating bacteria, they serves as adhesisn and are bound with high affinity by type IV pili. The eDNA stabilizes the biofilm, contribute in antibiotic resistance and cation gradients, genomic DNA release, as well as a nutrient source for starving bacteria after degradation by extracellular deoxyribonucleases (Wei *et al.*, 2013). In addition, eDNA forms bundles with F-actin liberated from necrotic neutrophils and can initiate biofilm development (Parks *et al.*, 2009).

A cellular appendages such as pili, fimbriae and flagella are also consider as a matrix components. These structures are important for adhesion and/or motility, as well as structural support (Mann *et al.*, 2012).

Biofilm formation

Biofilm formation has been divided into five stages:

1. Attachment and reversible adhesion- planktonic bacteria encounter a surface and with the use of flagella and type IV pili attach loosely and reversibly to it. This is primarily influenced by electrostatic forces, van der Waals forces and Brownian motion. At this stage, the bacteria are still susceptible to antibiotics, in accordance with the success of perioperative antibiotic prophylaxis (Høiby *et al.*, 2010; Wagner *et al.*, 2008).
2. Irreversible adhesion- within a few hours from the initial adhesion, the bacteria starts its multiplication and extracellular metabolites production (EPS and proteins), which lead to matrix formation bind to the surface irreversibly. The main role is played by non-specific interactions (e.g. hydrophobic, hydrogen and ionic bonds) and specific interactions (e.g. adhesins) (Høiby *et al.*, 2010).
3. Maturation I – biofilm grows in thickness (up to 50μm) and mushroom-like or tower-like structures are often observed (*in vitro* conditions). The type IV fimbriae play an important role here.
4. Maturation II- depends on the availability of nutrients, osmolarity, the oxygen concentration and pH within the biofilm. New proteins are synthetized, that were absent

in planktonic form, and alginate is overproduced. At this stage, the biofilm shows maximum tolerance/resistance to antibiotics (Høiby *et al.*, 2010).

5. Dispersion- local areas of the biofilm dissolve and cells can detach via active or mechanical (e.g. increased external fluid shear) means and then spread to another location where new biofilms may form. This liberation process may also be caused by bacteriophage activity within the biofilm (Høiby *et al.*, 2010).

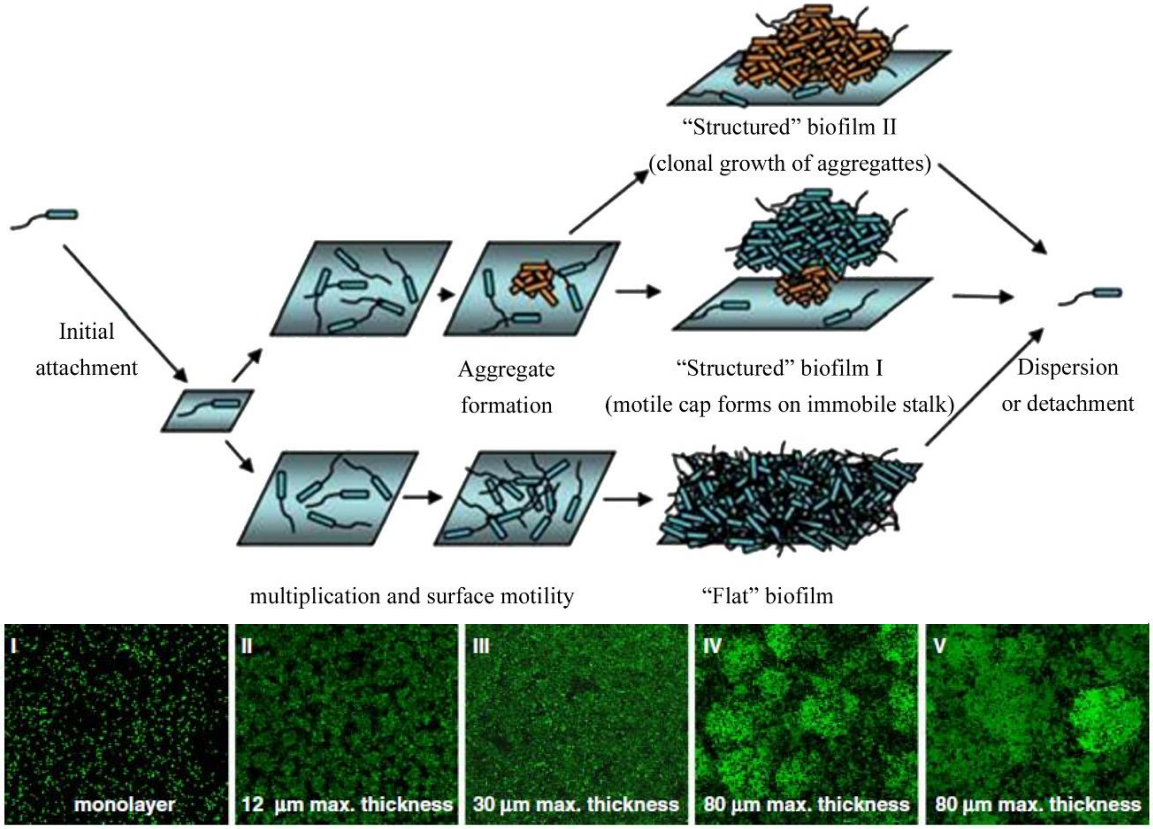


Figure 1.4: Representation of *P. aeruginosa* flat and structured biofilm development. Initial attachment involves adherence of free-swimming cells to the surface. In the case of a flat biofilm, cells continue to multiply and move on the surface, forming a confluent, flat mat of cells. In the case of structured biofilms, there are two alternative routes to their formation. The first is 'Structured Biofilm I'. Here, orange cells represent the immobile 'stalks' of structured biofilms, while light blue cells represent the motile subpopulation. The second is 'Structured Biofilm II'. In this case small cell aggregates grow clonally, forming large cell aggregates consisting of cells primarily derived from cells in the small cell aggregates. This is indicated by the large aggregate of orange cells in the figure. The cells in aggregate are progeny derived from the initial small aggregates. Finally, cells can actively leave the biofilm to reinitiate the cycle in a process called dispersion or detachment (Kirisits *et al.*, 2006). Below the scheme are a two-dimensional projection of the three-dimensional image shown as a top-down view for 2 h (I), 1 day (II), 2 days (III), 3 days (IV), and 4 days (V) after inoculation. Maximal thickness of the biofilm is reported for each time period. Biofilm images were collected by a Leica confocal scanning light microscopy using a 488-nm wavelength. Images are at ×400 and represent a 250×250-µm field (Wagner *et al.*, 2008).

The impact of biofilms in industry

In industry, biofilms have become problematic in a wide range of food industries, such as brewing (Flemming *et al.*, 2009), seafood processing (Shikongo-Nababi, 2011), dairy processing (Chmielewki *et al.*, 2003), poultry processing (Harvey *et al.*, 2007), meat processing (Sofos *et al.*, 2010) and equipment contamination (e.g. pipelines) (Stewart *et al.*, 2008). In medicine, up to 80% of persistent bacterial infections were found to be biofilm associated (Janssens *et al.*, 2008; Srey *et al.*, 2013; Blair *et al.*, 2014). It occurs on natural surfaces such as teeth, heart valves (endocarditis), in the lungs of CF patients causing chronic bronchopneumonia, in the middle ear in patients with chronic and secretory otitis media, in chronic rhinosinusitis, in chronic osteomyelitis and prosthetic joint infections, in intravenous catheters and stents and in chronic wounds (Høiby *et al.*, 2011).

Biofilms are highly resistant to antibiotics and chemical disinfectants, as well as the host immune system (Khalifa *et al.*, 2011). There are several factors of unique biofilm defense: 1) size exclusion and delay of antibiotic diffusion through the pores of EPS matrix (e.g. aminoglycosides); 2) synthesis of periplasmic glucans that sequester cationic antibiotics (Ghafoor *et al.*, 2011); 3) an altered metabolic activity correlated to cell position in the heterogeneous biofilm structure (e.g. β−lactams, ciprofloxacin) (Klausen *et al.*, 2003); 4) unequal cells susceptibility to antimicrobial agents, presence of subpopulations with a resistant phenotype: persisters in dormant state and phenotypic variants originate from mutations and genetic rearrangements, that can grow normally in the presence of elevated levels of antibiotics (Harmsen *et al.*, 2010).

1.1.3. Antibiotic resistance: bad bugs, no drugs: No ESKAPE!

The arrival of the antibiotic era largely palliated the previously fatal outcomes of acute infections in patients. Standard antibiotic therapy in use against *P. aeruginosa* infections include a number of β-lactams such as cephalosporins (e.g. ceftazidime), carbapenems (e.g. imipenem, meropenem), monobactams (e.g. aztreonam), and penicillin (e.g. piperacillin). Other major classes of antibiotics used for treatment comprise aminoglycosides (e.g. gentamicin, tobramycin, amikacin), quinolones (e.g. ciprofloxacin), and polymyxins (e.g. colistin, polymyxin B) (Taylor *et al.*, 2014). However, the inappropriate antibiotic treatments and the development or acquisition of multidrug-resistant strains lead to resistant chronic *P. aeruginosa* infections (Kang *et al.*, 2005, Lodise *et al.*, 2007, Micek *et al.*, 2005).

Antibiotic resistance results from several independent mechanisms. They can be divided into three models: intrinsic, acquired and adaptive resistance. First, include a reduced permeability of the outer membrane by presence of small size porins (e.g. vancomycin resistance) (Nicas *et al.*, 1983; Hancock, 1998; Blair, 2014). Second, active removal of antibiotics by efflux pumps and degradation through enzymatic modifications. These factors make *P. aeruginosa* resistant to penicillin G, aminopenicillins, macrolides, cephalosporins I and II generation, tetracyclines, sulfonamides, and chloramphenicol. Acquired resistance may be of endogenous (mutational variation) or exogenous origin (horizontal gene transfer by e.g. by conjugation, plasmids, transposons, integrons). The endogenous factors include: cephalosporinases AmpC derepression, overexpression of efflux pumps, porin OprD repression, mutations in gyrase and topoisomerase IV (Blair, 2014). Third, resistance can develop during treatment (e.g. repeated exposure to sublethal concentrations of antibiotics). In this case, the bacteria adapt themselves to survive by undergoing chromosomal mutations in genes. Other mechanisms of resistance, such as ESBL (β-lactamase extended substrate spectrum) and MBL (metallo-β-lactamase) and aminoglycosides modifying enzymes arise from horizontal gene transfer (Wolska *et al.*, 2013). *P. aeruginosa* resistance mechanisms to antibiotics are presented in Table 1.3.

Finally, persister cells of *P. aeruginosa* exhibit multidrug tolerance, which, in contrast to resistance, results from dormancy rather than mutations. The cells can survive but not grow in the presence of antibiotics. When antibiotic pressure drops, the persister cells can restore the population, with the same antibiotic susceptibility features as original one with additional small pool of persister cells (Fauvart *et al.*, 2011). In CF environment resistance of chronic *P. aeruginosa* infections largely results from multidrug-tolerant subpopulation of persister cells (Lewis, 2007; Fauvart *et al.*, 2011). Moreover, high persister (hip) mutants have been isolated from CF patients, the appeared probably in response to the periodical administration of antibiotics (Lewis, 2007, Mulcahy *et al.*, 2010).

Table 1.3: Resistance mechanisms (Gellatly *et al.*, 2013; Strateva *et al.*, 2009).

Mechanism	Resistance class	Examples
Outer membrane impermeability	Intrinsic	Outer membrane channels OprF, OprD, OprB (carbapenems, aminoglycosides, quinolones) EPS
Efflux pumps	Intrinsic	MexAB–OprM (Quinolones, macrolides, tetracyclines, lincomycin, chloramphenicol, novobiocin, b-lactams except imipenem) MexCD–OprJ (Quinolones, macrolides, tetracyclines, lincomycin, chloramphenicol, novobiocin, penicillins except carbenicillin and sulbenicillin, cefepime, cefpirome, meropenem) MexEF–OprN (Fluoroquinolones, carbapenems) MexXY–OprM (Quinolones, macrolides, tetracyclines, lincomycin, chloramphenicol, aminoglycosides, penicillins except carbenicillin and sulbenicillin, cefepime, cefpirome, meropenem)
β-lactamases	Intrinsic	AmpC (penicillins)
Targeted mutation	Acquired	DNA gyrase, DNA topoisomerase IV, 16s rRNA methylation (quinolones)
Horizontal transfer	Acquired	Metallo-β-lactamases, ESBLs (penicillins, cephalosporins, carbapenems)
Membrane changes	Adaptive	Lipid A modification (aminoglycosides, polymyxins) AmpC upregulation (penicillins)

1.2. Bacteriophages

1.2.1. Short history of phage research

In 1896 British bacteriologist Ernest Hanbury Hankin observed during a cholera epidemic in India, that people drinking water from the Ganges and Yamuna rivers do not get sick. Examination of the water samples revealed the existence of "substances" with antibacterial properties, that prevented the spread of the epidemic. Several years later, two scientists, British bacteriologist Frederick Twort (1915) and French-Canadian microbiologist Félix d'Herelle (1917) independently observed virus-associated cell lysis of *Micrococcus* (currently *Staphylococcus aureus*) and *Shigella* sp. D'Herelle was also the first scientist who applied the term "*bacteriophage*" as "an invisible, antagonistic microbe of the dysentery bacillus" (Adams, 1959; Kutter, 2005). In the 1920s and 1930s Lilly and Squib started to work on bacteriophage preparations for the treatment of *Staphylococcus* infections. In 1923, the Eliava Institute was opened in Tbilisi, Georgia, which was the phage therapy center behind the 'Iron Curtain' for almost 70 years. To date, it is the most well known phage therapy institute. However, a number of factors including antibiotic discovery caused the decreased use of bacteriophages for medical applications for the next decades (Sabouri Ghannad *et al.*, 2012).

1.2.2. Bacteriophage classification

The name "*bacteriophage*" is derived from the Greek *bacterion* (bakteria, rod, stick) and *phagein* (eat, devour). Bacteriophages (phages) are obligatory intracellular bacterial parasites. They are highly diversified and abundant in biosphere, outnumbering prokaryotes approximately ten fold (Donlan, 2009; Kohanski *et al.*, 2010). The bacterial host is required for viral progeny reproduction (Donlan, 2009). This situation leads to an continuous arms race and coevolution of host and its infecting virus, resulting in vast variety of phage species with different mechanisms of host infection (Le *et al.*, 2014). Most of phages are highly specific and have a narrow host-range, targeting only few strains within a species (Stewart *et al.*, 2008).

Phages are classified based on (i) infection cycle (lytic or temperate), (ii) morphology (filamentous or icosahedral phages with/without tails, phages with a lipid-containing envelope or contain lipids in the particle shell) (Fig. 1.5), (iii) genome

composition (dsDNA, ssDNA, ssRNA, dsRNA or segmented) and (iv) genome organization (subfamilies & genera).

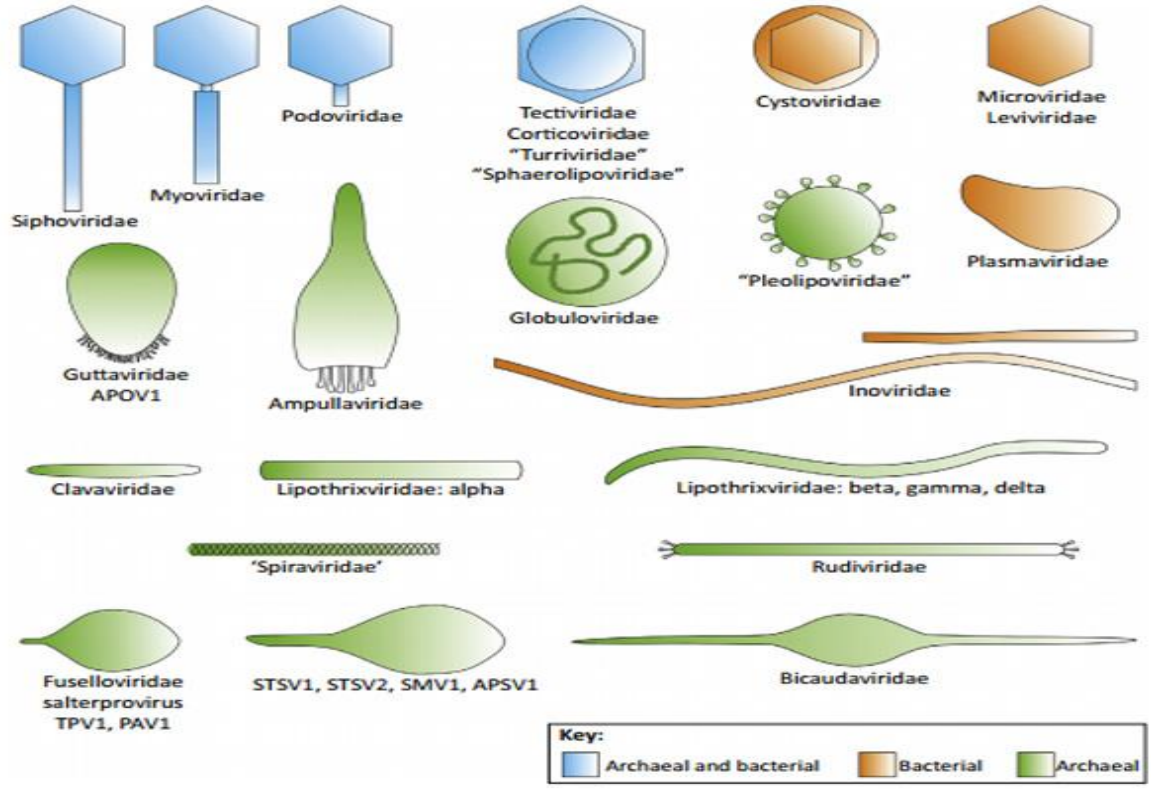


Figure 1.5: Classification of microbial viruses based on morphology. Virion morphotypes of prokaryotic and archaeal viruses. Abbreviations: APOV1, *Aeropyrum pernix* ovoid virus 1; APSV1, *Aeropyrum pernix* spindle-shaped virus 1; PAV1, *Pyrococcus abyssi* virus 1; SMV1, *Sulfolobus monoca*. Viruses infecting archaea and bacteria are marked blue, only bacteria orange and only archaea green (Pietilä *et al.*, 2014).

Phages have evolved at least three distinct types of infection cycles: strictly lytic, lysogenic or pseudolysogenic. The lytic cycle leads to cell lysis and obviously death of the bacterial host. It has several steps: 1) phage attachment to specific receptor on the host cell surface (e.g. LPS, OM porins, flagella or pili) with/without production of EPS degrading enzymes. This step determines phage specificity and its host range; 2) injection of genomic DNA into the bacterial cell; 3) transcription and translation of early, middle and late phage genes by hijacking the host synthesis machinery; 4) formation of new virion particles; 5) disruption of the bacterial cell wall by peptidoglycan degrading enzymes, called endolysins, and release of the progeny (Donlan, 2009). In general terms, temperate phages have genetic switches, which allows them to determine if lytic infection occurs or expression of phage genome will shutoff and the genome will frequently be incorporate in the bacterial genome (prophage state). The switch occurs at random or can be triggered by environmental stresses when the survival of the host bacterium is not assured (Clokie *et al.*, 2011). Finally, when bacteria are starved or experience suboptimal growth, and

exhibit low metabolism and reduced growth, pseudolysogeny may occur. It is a stalled phage development in a host cell without multiplication of the phage genome or its replication, a survival mechanism in hostile environment. Phage genome is present within the host cells in an unstable, inactive state and may stay present unexpressed for a long period. When nutrients become again available, the pseudolysogenic state can be reversed and gene expression can take place (Los *et al.*, 2012).

Approximately 96% of phages have a complex morphology structure (icosahedral head and tail), while ~ 4% have isometric, helical or pleomorphic structure (Fig. 1.6) (Matsuzaki *et al*., 1998).

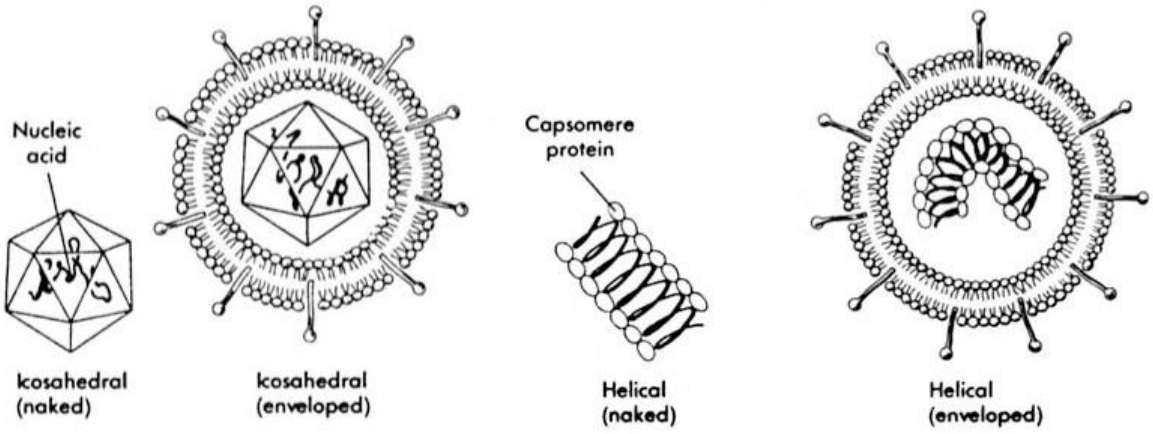


Figure 1.6: Phage capsid morphology. From left to right: icosahedral (naked), icosahedral (enveloped), helical (naked) and helical (enveloped) structures (Towar K, 2011).

On the basis of the tail morphology, order *Caudovirales* is divided into three families (Fig. 1.7): 1) *Myoviridae* with contractile tails consisting of a sheath and a central tube; 2) *Siphoviridae* with long, noncontractile tails, and 3) *Podoviridae* with short tails (Kurtböke *et al.*, 2012). The viral capsid holds packed genetic material and proteins, while tail and its fibers are responsible for the host cell recognition, attachment, cell envelope penetration and delivery of the phage DNA into the host cell cytoplasm (Leiman *et al.*, 2004 and 2012; Davidson *et al.*, 2012).

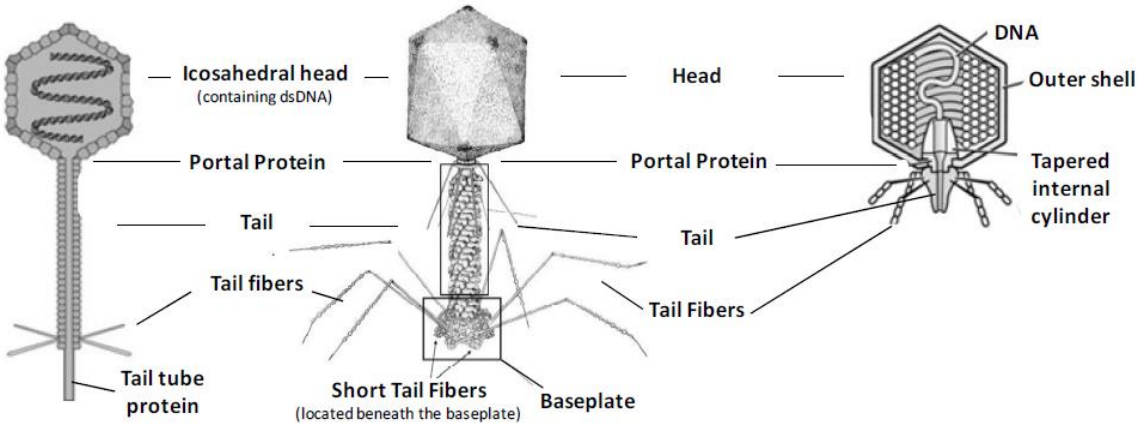


Figure 1.7: Morphology of phages belonging to order *Caudovirales.* From left to right, schematic drawings of a *Siphoviridae* family (icosahedral head and long noncontractile tail, e.g. λ phage), *Myoviridae* family (icosahedral head and long contractile tail, e.g. T4 phage) and *Podovidae* family (icosahedral head and short tail, e.g. T7 phage) are shown (Serwer *et al.*, 2008).

The phage virion may consist of different genetic material (ss/dsDNA or ss/ds RNA), enclosed within a protein or lipoprotein coat. The size of the phage genomes vary from a few kilobase pairs (kbp) to several hundred. In last decades the number of sequenced phage genomes increased as well as the number of unknown ORFs (Open Reading Frames) without or low sequence homology. These genes are called 'viral dark matter' and are a vast reservoir of genetic information (Rohwer *et al.*, 2011). Many of them are non-essential, however some might have important function and influence on e.g. phage fitness in specific environments (Hargreaves *et al.*, 2014).

1.2.2.1. *Pseudomonas* bacteriophages

In 1964 Lindberg assembled a *P. aeruginosa* typing set, among which Ackermann *et al.* (1988) found eight different morphotypes (Fig. 1.7).

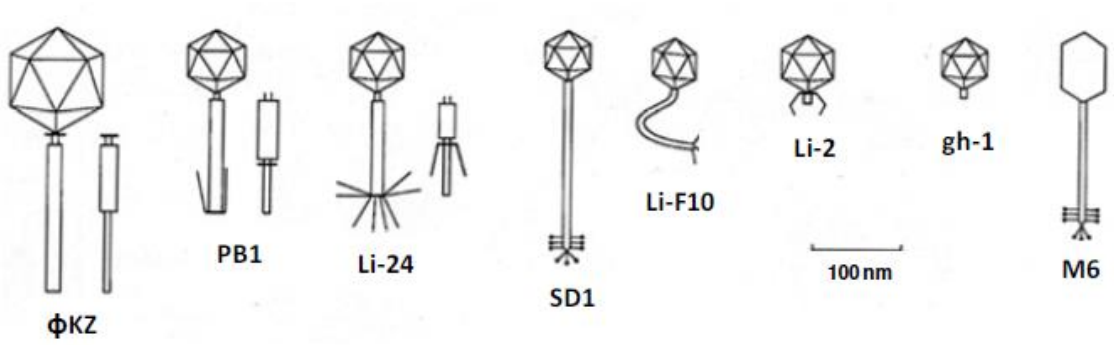


Figure 1.7: Scale drawing of the different morphotypes of the phages belonging the Lindberg typing set (Ackermann *et al.*, 1988).

Later, Ackermann also examined after negative staining over 5500 phages in the electron microscope among which 511 infected *Pseudomonas* genus (Ackermann,

2007). More than 97% of this pool were tailed phages, while only 3% were polyhedral, filamentous or pleomorphic. Moreover, the majority belong to the *Siphoviridae* family (47%). In recent years, formal taxonomy has relied heavily on the availability of complete genome sequences of these phages, which has led to the creation of several new genera and subfamilies based on genome organization. Interestingly, this molecular classification still corresponds well to the original morphotype classification (www.ICTVonline.org).

Among the *Myoviridae* family six types infecting *Pseudomonas* can be distinguished: *P2virus*, a genus of the *Tevenvirinae* subfamily and several independent genera: *Kpp20virus, Pakpunavirus, Pbunavirus* and *Phikzvirus* (according to International Committee on Taxonomy of Viruses, ICTV). The most well know and studied is phage φKZ, originally isolated in 1975 from sewage sample in Alma Ata, Kazakhstan. It represents a giant and broad-range genus of *Phikzvirus.* Its virion has large head of 145 nm diameter and long contractile tail of 200 nm, surrounded by a loose fibrous network (Ackermann *et al.*, 1988; Mesyanzhinov *et al.*, 2002). Its capsid contains no less than 50 different proteins and holds a 280 kb dsDNA, circularly permuted and terminal redundant genome, spooled around protein inner body (Thomas *et al.*, 2012). The majority of predicted ORFs (75% out of 369 predicted ORFs) share no homology with other proteins in NCBI database, which results in low relatedness to other phages (Krylov *et al.*, 2007), except *Pseudomonas* phage PA7.

All *Pseudomonas* phages within the *Siphoviridae* family described to date are temperate and represent transposable phages, such as phage B3, D3112 and D3. The ICTV database lists five types of *Pseudomonas Siphoviridae: D3112virus*, *D3virus, Septima3virus* and *Yuavirus,* which are often isolates from clinical environments (Krylov *et al.*, 2007, Zegans *et al.*, 2009). Currently, it is the most studied group among *Pseudomonas* phages. Phage D3112 is one of the representative siphoviruses. It has a number of important characteristic features: 1) rare palindrome GTCGAC and CTGCAG in its genomes; 2) absence of autoxotrophic mutants among lysogens; 3) many sites for integration into a bacterial chromosome and 4) modular genomes (around 37 kbp, 50 - 60 predicted ORFs) with the ability to exchange modules by recombination (Krylov *et al.*, 2007).

Currently, the ICTV database distinguishes five types of *Podoviridae* phages infecting *Pseudomonas: Phikmvvirus, T7virus, F116virus, Lit1virus, Luz24virus*. They are divided based on genome organization and behavior during infection. The lytic *Pseudomonas* phage gh-1, isolated in 1966, is a representative of *T7virus,* that forms large

clear plaques and requires O-antigen on the bacterial surface as a receptor. It has a 37 kbp linear genome and 42 predicted ORFs. Among the *T7virus* genus there is strong correlation in genome organization and conserved replication strategy (Kovalyova *et al.*, 2003;Rehm, 2009).

1.2.3. Bacteriophage - based therapeutics

Phages exhibit several advantages for application as antimicrobials: 1) high diversity and abundance; 2) bacteriolytic activity; 3) high specificity to one genus or species; 4) they can work in synergy with other phages with different lysis mechanisms and/or bacterial surface receptors in cocktail preparations, as well as with antibiotics; 5) self-dosing therapy; 6) low costs preparations (Hyman *et al.*, 2012; Bikard *et al.*, 2014; Brown-Jaque *et al.*, 2015; Roach *et al*, 2016). However, also number of key disadvantages are present, including the presence of toxins or integrases encoded by temperate phages, and the potential for horizontal gene transfer between phages and host (e.g. transfer of virulence genes) via specialized or generalized transduction (Brown-Jaque *et al.*, 2015; Roach *et al*, 2016). In the last decades, the absence of preclinical and clinical data led to prohibition of phage therapy around the world, except in Georgia, Russia and Poland (Chanishvili, 2012). Currently, the commercial interest has been recovered by small pharmaceutical companies focusing on phages (Table 1.4) and phage-derived proteins (this section will be discussed in details in further sections).

Table 1.4: Current phage therapeutics development for human diseases treatment (Cooper *et al.*, 2016).

Company	Product	Type	Target	Application	Company website
Intralytix	ShigActive	Phage	*Shigella*	Ingested	http://www.intralytix.com/index.php?page=prod
AmpliPhi	AmpliPhage-001 AmpliPhage-002 AmpliPhage-004	Phage	*P. aeruginosa* *S. aureus* *C. difficile*	-* Topical -	http://www.ampliphibio.com/product-pipeline.html
Technophage T	TP-102 TP-122 TP-132 TP-107	Phage	-* -* -* -*	Ulcers Respiratory -* Topical	http://www.technophage.pt/index.php/r-d/productpipeline
Pherecydes Pharma	PP021 PP1131PP1231 PP2351	Phage	*E. coli* *Pseudomonas* *Staphylococcus*	Burn and Skin Burn, Skin, and Respiratory tract infection Bone, Joint, and Prosthesis	http://www.pherecydes-pharma.com/pipeline.html

(*) information not available

Within the field of food and agriculture industry, several phage preparations have already been approved in USA as a food protecting agents (LISTEX™, SALMONELEX™, Listshield™, Ecoshield™, SalmoFresh™) (Rodríguez-Rubio *et al.*, 2015).

1.3. Phage enzymes with antimicrobial properties & applications

1.3.1. Peptidoglycan degrading enzymes

1.3.1.1. Bacterial cell wall structure

The bacterial cell wall consist of several layers and its rigidity depends on an internal pressure, correlated with the osmotic forces (Schlegel, 2008). Bacteria possess a peptidoglycan (PG) layer, also known as a murein sacculus (Slonczewski *et al.*, 2011; Hyman, 2012). Murein is a heteropolymer composed of alternating *N*-acetylglucosamine (GlcNAc; an amino sugar derivative) and *N*-acetylmuramic acid (MurNAc; glucosamine with a lactic acid group) chains, which are linked by β-1,4-glycosidic bonds. These glycan chains are connected by short peptides consisting of four to six amino acids, such as L-alanine, D-glutamic acid, *meso*-diaminopimelic acid (mDAP) or L-lysine and D-alanine (Schlegel, 2008). The MurNAc lactic acid residue, L-alanine, is the link between the amino acid of the stem and the sugar backbone (Schleifer *et al.*, 1972).

In Gram-positive bacteria murein layer is thick (~40 layers, 20 – 80 nm) and comprises 30 - 70% of a cell wall weight. Often instead of mDAP, LL-diaminopimelic acid or L-lysine occur in the third amino acid position. Another characteristic feature is the presence of teichoic acids, composed of 8 - 50 glycerol or ribitol molecules conntected by phosphodiester bridges. Teichoic acids might be attached to peptidoglycan layer by phosphates in amide-like bond (Fig. 1.8) (Schlgel, 2008). In contrast, Gram-negative bacteria have a thinner peptidoglycan layer (1 to 3 layers, 10 nm) which comprises less than 10% of cell wall weight. The murein lacks a lysine residue at the third amino acid position of the dipeptide and instead has mDAP, interlinked with a D-alanine of an adjacent peptide stem from another glycan chain. Importantly, the cell wall has an additional outer membrane rich in lipopolysaccharides, lipoproteins and phospholipids (Fig. 1.8) (Madigan *et al.*, 2008; Schlgel, 2008; Cornelissen *et al.*, 2012).

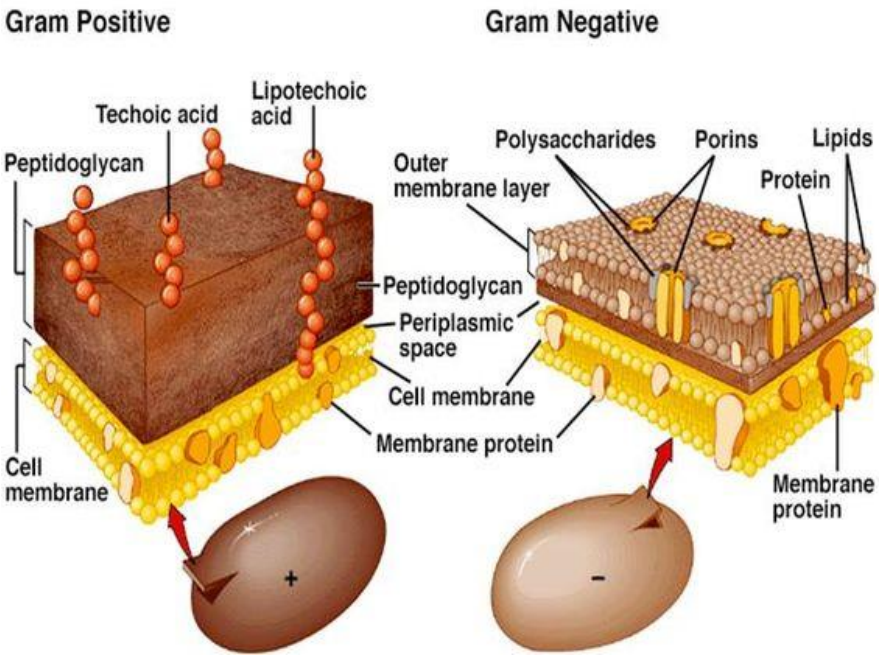


Figure 1.8: The Gram-positive and Gram-negative bacteria cell wall comparison. From left to right: Gram-positive bacteria with an inner cell membrane and thick peptidoglycan layer; Gram-negative bacteria with inner cell membrane, thin peptidoglycan layer and outer membrane layer with surface proteins, porins and lipids (Aryal, 2015).

1.3.1.2. General characteristics of PG degrading enzymes

Bacteriophage-encoded peptidoglycan degrading enzymes are bacterial cell wall degrading proteins which play a key role in the lytic replication cycle (Fig. 1.9). First, virion-associated lysins locally degrade peptidoglycan and allow the virus to penetrate through the cell wall and inject its genome. Further, in the late stage of the cycle, genes responsible for a two component holin-endolysin mechanism (specific to *Caudovirales*) and spannins are overexpressed and phage progeny are released by bursting the cell envelope (Stone, 2002).

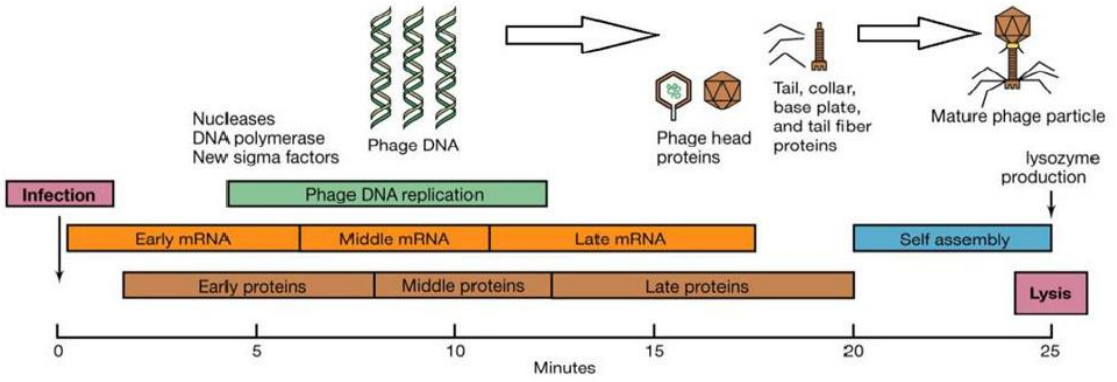


Figure 1.9: Timeline of the T4 phage replication pathway. The chronological order of T4 lytic cycle and expression of early, middle and late genes. Early genes include enzymes modifying the host's DNA replication and transcription. Middle genes are responsible for phage DNA replication and late genes activation. Late genes play key role in production and assembly of structural proteins followed by self-assembly and progeny release (Madigan *et al.*, 2008).

Peptidoglycan degrading enzymes are considered to be efficient antimicrobials against most common pathogens such as: *Staphylococcus* sp., *Streptococcus pneumoniae, Enterococcus faecium, Bacillus anthracis, Clostridum* sp., *Klebsiella pneumonie* and *P. aeruginosa* (Stone, 2002).

1.3.1.3. Structure of PG degrading enzymes

There are three types of phage bacteriolytic PG degrading enzymes: endolysins, virion-associated PG hydrolases (VAPGHs) (Fig. 1.10) and spannins. They have similar mode of action, yet differ in structure and time of action during phage lytic cycle.

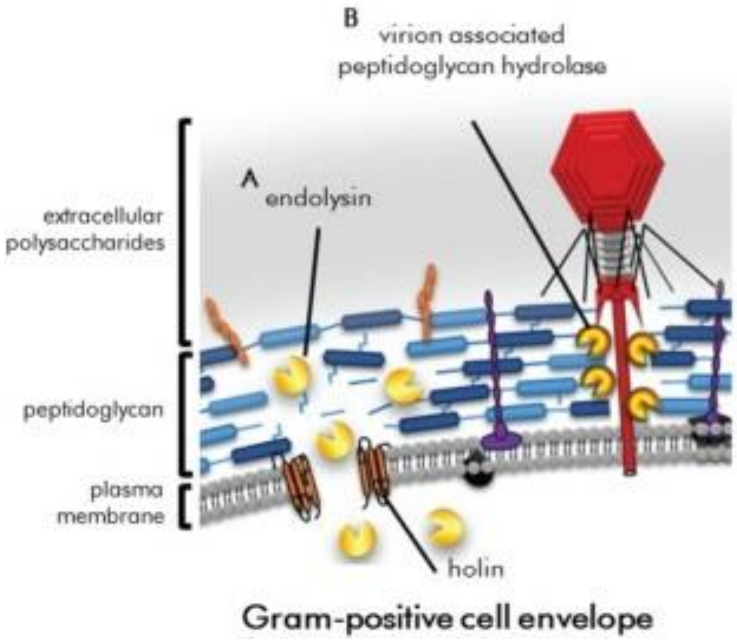


Figure 1.10: Schematic representation of PG layer degradation by application of endolysins, VAPGHs and holins. The peptidoglycan bonds are cleaved, which leads to cell wall instability and the high internal pressure of cell protrudes the plasma membrane resulting in osmolysis (lysis from without) (Roach *et al*., 2015).

VAPGHs are tail-associated muralytic enzymes, important at the beginning of the replication cycle. They are located beneath the baseplate and assist in phage genome injection by localized PG degradation (Rodríguez-Rubio *et al.,* 2015). Their activity is rather superficial since the cell wall stability has to be ensured for the bacterial host survival (Rodríguez-Rubio *et al.*, 2013). They are present in Gram-positive and Gram-negative phages, however not all are required for successful genomic transfer (Roach *et al*., 2015). VAPGHs consist one or two enzymatically active domains (EADs), that can cleave PG on several positions, either in the glycan moieties (by glycosylases) or the peptide stem (by amidases/peptidases) (Fenton *et al.*, 2010; Fischetti, 2010; Nelson *et al.*, 2012).

In contrast, endolysins act from within. In the final step of replication cycle they are accumulated in the cytosol to lyse the bacterial cell wall (Fenton *et al.*, 2010; Nelson *et al.*, 2001). However, the enzymes have several strategies to cross the cell membrane (Fig. 1.11). Often, they require assistance of secondary proteins, called holins. These small hydrophobic proteins with transmembrane domains are able to permeabilize the cytoplasmic membrane by forming pores in a concentration-dependent manner (Wang *et al.*, 2000). This results in release of the endolysins to the periplasm and PG layer destabilization (Wang, 2006).

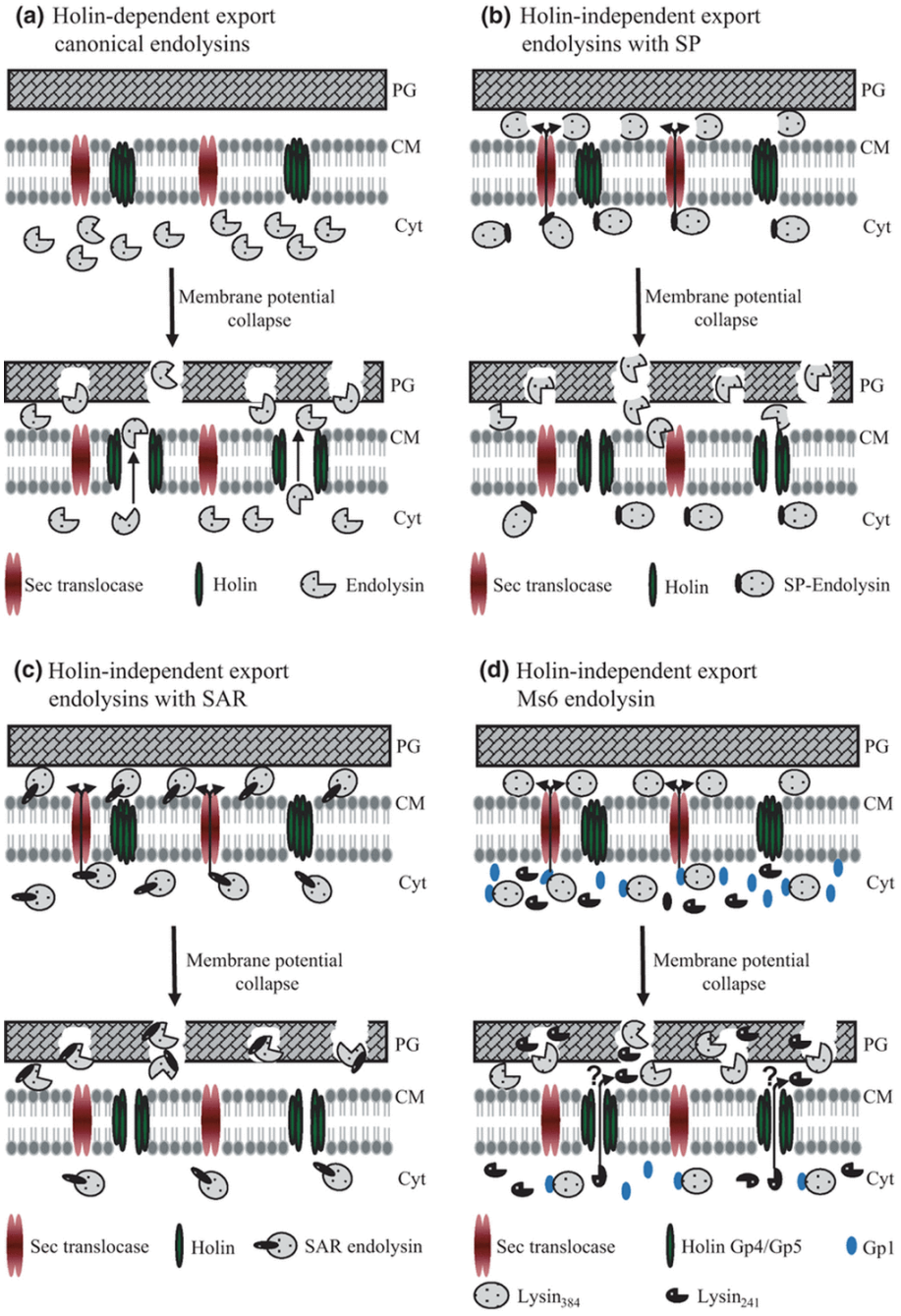


Figure 1.11: Schematic representation of different models for export and activation of phage endolysins. (a) In phages such as λ the export of the active endolysin to the cell wall is through the holin

pores. Holin-independent, Sec-mediated export of endolysins is observed in: (b) phages producing endolysins with typical signal peptides (SP), such as oenophage fOg44; (c) in phages synthesizing signal-arrestrelease (SAR) endolysins, as observed in coliphage P1; and (d) in mycobacteriophage Ms6, where export of the full-length endolysin (Lysin384) is assisted by the chaperone Gp1. When endolysins are exported through the Sec translocase, they are maintained in an inactive state in the cell wall compartment until holins dissipate the membrane pmf. The endolysin activation after pmf collapse is schematically represented by the change of the enzyme spherical configuration to a 'pacman' shape. Lysin241 in (d) is an N-terminally truncated version of the Ms6 endolysin produced by internal translation signals in the same reading frame. (?) indicates that export of this shorter version to the extracytoplasmatic environment is not known. (PG) peptidoglycan, (CM) cytoplasmic membrane, (Cyt) cytoplasm (Catalão *et al.*, 2013).

Despite their conserved biological function, endolysins are heterogeneous in structure and enzymatic activity. This divergence is directly related to the Gram-positive and Gram-negative bacterial cell wall structure. They can be globular or modular enzymes. The globular enzymes are mainly associated with phages infecting Gram-negative bacteria and usually contain a single EAD that cleaves specific bonds within PG (Briers *et al.*, 2007). Modular enzymes are usually associated with phages of Gram-positive bacteria, and contain an N-terminal EAD and a C-terminal cell wall-binding domain (CBD) (Fischetti, 2010; Nelson *et al.*, 2012; Walmagh *et al.*, 2012) with few exceptions. The CBD is responsible for high specificity through recognition and noncovalent binding to epitopes on the bacterial cell wall, such as carbohydrates, teichoic acids or peptides (Fenton *et al.*, 2010). Also endolysins with two EADs and absence of CBD have been reported. Sometimes an additional flexible interdomain linker sequence can be present. It provides additional flexibility in the structure, which allows domains to work autonomously (Fenton *et al.*, 2010; Rajaure *et al.*, 2015).

The CBDs prevent the diffusion of enzyme to intact with adjacent cells, by tightly binding the cell wall, even after cell lysis. However, the Gram-negative infecting phages have to overcome a much thinner PG layer but with the presence of the OM, the collateral damage by endolysin diffusion after lysis is prevented, which eliminates the need for CBDs. This is why most of those phages possess smaller, globular endolysins (Briers *et al.*, 2015).

In phages infecting Gram-negative bacteria the third class of lysis proteins called 'spanins' are present, responsible for OM disruption, possibly due to fusion with the cytoplasmic membrane (Young, 2014; Briers *et al.*, 2007).

1.3.1.4. Mode of action of PG degrading enzymes

The conserved overall composition and architecture of the bacterial PG permits the cleavage of only limited types of covalent bonds. However, there is a clear discrepancy between Gram-positive and Gram-negative bacteria. The first lack an OM and have a diverse PG composition, which results in over 100 different known chemotypes.

The GlcNAc and MurNAc backbone is conserved, with only rare examples of O-acetylation, N-glycosylation or N-deacetylation (e.g. *Bacillus* sp., *Micrococcus luteus* or *Mycobacterium smegmatis*), while the peptide moiety (amino acids composition and interpeptide cross bridging) is a source of PG diversity (Schleifer *et al.*, 1972; Vollmer, 2008). This is why phages infecting Gram-positive bacteria have to be armed with highly diverse and specific endolysins. In case of Gram-negative the diversity and specificity is lower, due to reduced selective pressure of PG in the presence of OM, low variation in the polysaccharide backbone, the interpeptide bridges and crosslinks. As a result, the endolysins of phages infecting Gram-negative bacteria have broad enzymatic activity (Walmagh *et al.*, 2013; Schleifer *et al.*, 1972).

The EADs comprise the catalytic activity of the endolysins and VAPGHs. Depending which specific bond is cleaved, they are classified into six groups (Fig. 1.12): (1) *N*-acetyl-β-D-muramidases (similar to lysozyme) cleave the *N*-acetylmuramoyl-β-1,4-*N*-acetylglucosamine bond, one of the two alternating glycosidic bonds of glycan strand. (2) *N*-acetyl-β-D-glucosaminidases hydrolyze the other glycosidic bond (*N*-acetylglucosaminyl-β-1,4-*N*-acetylmuramine) between the PG sugars. Both types of enzymes are hydrolases, requiring water to break the bond. However, there also exist lytic transglycosylases, which cleave the β-1,4 linkages between MurNAc and GlcNac resulting in the formation of a 1,6-anhydromuramoyl product, without water presence for reaction catalysis. This enzymes catalyzes an intramolecular reaction, in which the neighboring sugar moieties serve as an electron acceptors and destroy the covalent bond between them. (3) *N*-acetylmuramoyl-L-alanine amidases that cleave the amide bond between sugar (glycan, MurNAc) and the peptide (L-alanine). This allows for disengagement of glycan polymer from the stem peptide and PG destabilization. Further are endopeptidases, divided in two classes: (4) L-alanoyl-D-glutamate endopeptidases and (5) interpeptide bridge-specific endopeptidases. Both types cleave the peptide bonds between two amino acids in the interpeptide bridges and crosslinks. Finally, (6) γ-D-glutaminyl-L-lysine endopeptidases cleave peptide bonds between amino acid residues within the peptide stem that interconnects the PG strands (Hyman, 2012; Fischetti, 2010; Fenton *et al.*, 2010; Loessner, 2005; Young, 2014; Briers *et al.*, 2007).

The most common, universal and broad spectrum are endolysins comprising amidase or muramidase activity, because they target the most conserved PG bonds. In contrast endopeptidases are more species-specific, hydrolyzing the most variable parts of PG, interpeptide bridges and crosslinks (Schmelcher *et al.*, 2012).

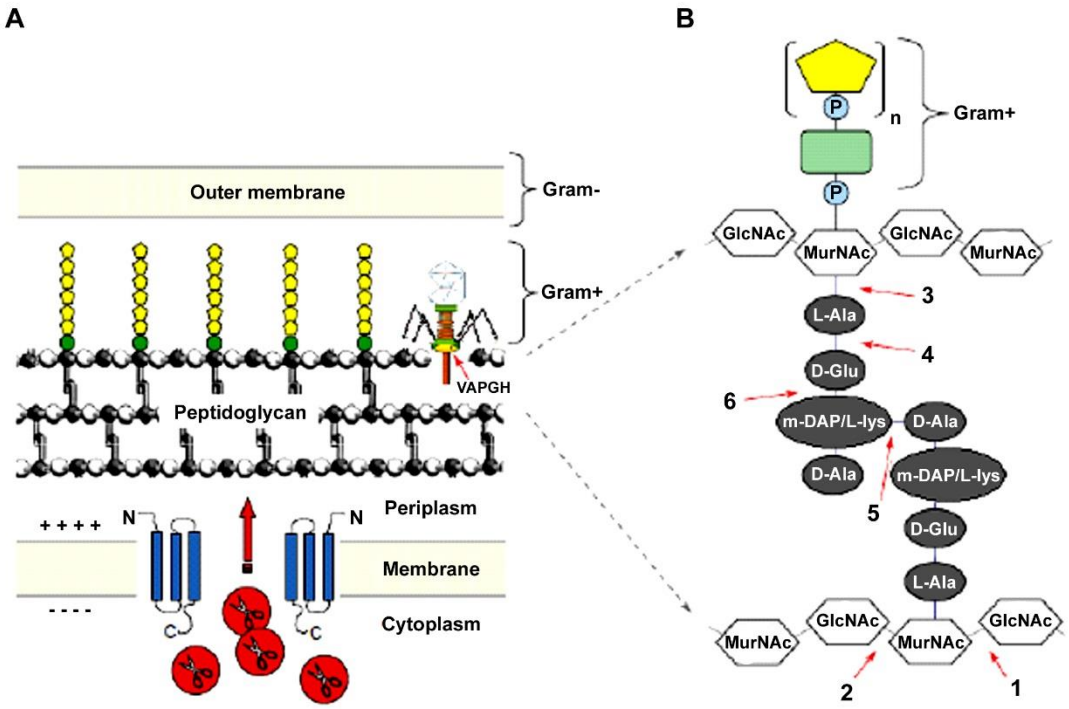


Figure 1.12: Mode of action of the PG-degrading enzymes. (A) Schematic representation of Gram-positive and Gram-negative bacteria cell wall structure. PG network marked as a grey balls, teichoic acids are marked as yellow chains. Holin proteins (blue) form membrane lesions by insertion of themselves into the cytoplasmic membrane and oligomerization. Next, the endolysins (red) pass through these pores to access the peptidoglycan. In case of Gram-positive cell walls with the thick and highly cross-linked peptidoglycan network lysis occur from the outside, while in Gram-negative cells lysis by free endolysin is prevented by the presence of the outer membrane. (B) PG structure magnification with possible PG hydrolase cleavage sites. The PG consists of alternating MurNAc and GlcNAc, which are crosslinked by peptide stems. The interpeptide bridge consists of a diamino acid (m-DAP) that is directly cross-linked to the terminal D-Ala of the opposite peptide chain. The bonds that can be cleaved by the EADs of endolysins are indicated. 1. *N*-acetyl-β-D-muramidase; 2. *N*-acetyl- β-D-glucosaminidase; 3. *N*-acetylmuramoyl-L-alanine amidase; 4. L-alanoyl-D-glutamate endopeptidase; 5. interpeptide bridge-specific endopeptidase; 6. γ-D-glutaminyl-L-lysine endopeptidase. Abbreviations: CCWP, carbohydrate cell wall polymer; GlcNAc, N-acetyl glucosamine; LU, linkage unit; m-DAP, meso-diaminopimelic acid; MurNAc, N-acetyl muramic acid; P, phosphate group (Loessner, 2005).

1.3.1.5. Bacterial resistance to PG degrading enzymes

To date, there are no reports describing endolysin resistance mechanisms or lost sensitivity to 'lysis from without'(Roach *et al.*, 2015). Several studies failed in resistance generation *in vitro* by repeated pathogen exposure to sub-lethal concentrations of endolysins or endolysin- sensitive strains mutagenesis (Schuch *et al.*, 2002). No resistance occurred or resistant mutants were not promoted, while in control groups, treated with only antibiotics, resistance increased 10- 100,000 fold.

However, the resistance to other types of PG hydrolases have been documented in case of human lysozyme with muramidase and cationic antimicrobial peptide activities. Resistance occurs due to modifications of peptidoglycan and/or cell wall linked components, such as teichoic acids (Davis *et al.*, 2011). *Staphylococcus simulans* secreted

lysostaphin, that binds to the cell wall envelope and cleaves the pentaglycine cross bridges of the *S. aureus* peptidoglycan matrix (Gründling *et al.*, 2006). Resistance occurs due to mutations and the subsequent formation of monoglycine cross-bridges, which eliminates the target site for the bacteriocins' catalytic site (Roach *et al.*, 2015). Pesticin secreted by *Yerisnia pestis* cells are protected by small immunity proteins, that binds and inactivates the enzyme (Lukacik *et al.*, 2012). Importantly, the extracellular mode of action of PG degrading enzymes hamper the majority of known antimicrobial resistance mechanisms, that act intracellularly (e.g. reduced membrane permeability, efflux pumps, inactivation by cytoplasmic enzymes). Moreover, the high abundance and variety of phages as well as protein engineering strategies provide a inexhaustible source of new enzymes (Roach *et al.*, 2015).

1.3.1.6. Potential applications

Phage endolysins and VAPGHs applied as a purified recombinant proteins can serve as efficient antimicrobial therapy alternative. Importantly, endolysins do not required holins to overcome membrane barrier and gain an access to PG layer. The external application of these enzymes to overcome common Gram-positive pathogenic bacteria such as streptococci, enterococci, *Clostridium difficile*, *Bacillus anthracis*, MRSA, etc. has been already extensively studied (Fenton *et al.*, 2010). Microbes can be lysed in few seconds after exogenous application of even a small amount of purified endolysins.

In the case of Gram-negative bacteria situation is more complicated, due to the presence of the OM, which serves as a very efficient shield against PG hydrolases. To circumvent this barrier several strategies were proposed. One idea is the combined administration with OM-permeabilizing agents, such as ethylenediaminetetraacetic acid (EDTA), detergents, etc. On the other hand, endolysins can be fused with protein transduction domains or peptides of non-endolysin origin to enable crossing the OM (Briers *et al.*, 2014). In 2014 Briers *et al.* presented the idea of protein- engineered endolysin (Artilysin®), where phage enzyme was fused with an OM permeabilizing peptide (OMP). This peptide can either be polycationic, in order to disrupt the stabilizing ionic bounds between divalent cations and the negatively charged phosphate groups, or hydrophobic, to interfere with the hydrophobic moieties present in the lipid A part of LPS (Briers *et al.*, 2014). Furthermore, Lukacik *et al.* (2012) developed hybrid protein, a three domain endolysin ('pesticin'), highly efficient against strains actively expressing FyuA, such as *Escherichia coli*, *Y. pestis* or *Yersinia pseudotuberculosis*. It consist the N-terminal

translocation domain, the FyuA OM transporter (a major virulence factor of *Y. pestis*) binding domain, and the C-terminal EAD, structurally similar to T4 lysozyme, that cleaved bonds in the PG layer (Lukacik *et al.*, 2012).

The advantages and limitations of PG degrading enzymes are summarized in Table 1.5.

Table 1.5: The benefits and challenges of PG degrading enzymes as antimicrobials (Rodríguez-Rubio *et al.*, 2016, Drulis-Kawa *et al.*, 2012, 2015).

Benefits	Challenges
high affinity	salinity, pH and temperature range limitations
fast kinetics outcompeting the host immune response	immunogenicity of proteins
synergy with antibiotics or other PG hydrolases with different catalytic specificity	release of immunomodulatory cell debris, such as teichoic acids, LPS etc.
lack of documented resistance development	lack of OM penetration ability (overcome by the Artilysin strategy)
single dose, continuous or repeated administration	protein dependent half- time
specificity with no effect on the commensal flora	.
noncorrosive and biodegradable	
reduction of antibiotic resistant mutants	
diversity in activity and structure	
high activity of low dosage	
suitable candidates for rational design and directed evolution	

The presented phage PG degrading enzymes opens the door to variety applications in biotechnology, food industry, agriculture and medicine (Fischetti, 2010). Table 1.6 shows several examples, where single or periodical phage enzymes administration overcome bacterial infections.

Table 1.6: The PG hydrolases potential application in industry and medicine

Application	Problem	Example	Reference
oral, nostril and topical administration	infections of eye, ear, skin and mucous membranes	(1) Streptoccocal C1 phage endolysin (PlyC, 500 U) administrated to the oral cavity impede group A streptococci colonization in upper respiratory track. Treatment fully eradicated bacterial infection up to 2h post treatment, while in control group infection occurred in 70% of mice. (2) single dose of phage-born endolysins reduced a *Streptococcus agalactiae* vaginal infections. (3) engineered VAPGH P128 (muralytic domain combined with specific cell wall targeting domain) was active against methicillin-resistant *S. aureus* in a nasal colonization model.	Nelson *et al.*, 2001 Cheng. *et al.*, 2005 Paul *et al.*, 2011
interperitoneal injection	systemic infections	PlyG endolysin of *B. anthracis* γ phage increased 70-80% mice survival in 72h long experiment, while control group died after 2h post infection.	Schuch *et al.*, 2002
vaccinations		The stimulation of the human adaptive immune system by viral nanoparticles (VNPs) vaccination employing the cell wall anchor of bacteriophage endolysin for surface display. Already recombinant endolysin Lyb5 with CBD domain from *Lactobacillus fermentum* temperate bacteriophage φPYB5 was successfully displayed on the surface of lactic acid bacteria, such as *Lactococcus lactis*, *Lactobacillus casei*, *Lb. brevis*, *Lb. plantarum*, *Lb. fermentum*, *Lb. delbrueckii*, *Lb. helveticus*, and *Streptococcus thermophiles*.	Proetzel. *et al*., 2012 Hu *et al.*, 2010
food preservation	food contamination by *S. aureus*, *Listeria monocytogenes, Salmonella sp.*, *Campylobacter sp.*, *E. coli*, *Clostridium sp.* and many other microbes	(1) direct endolysin application (e.g. LysH5 endolysin against *S. aureus* bovine and human strains in pasteurized milk); (2) diary fermentation strains engineering and direct endolysin secretion during fermentation process (e.g. Ply118 and Ply511 genes of *L. monocytogenes* phages introduced in a *Lactococcus lactis* fermentation strain); (3) developing of transgenic potato, expressing the T4 phage lysine active against infections of *Erwinia carotovora*; (4) detection of pathogenic strains by use of coated magnetic beads (e.g. CBD molecules fused with fluorescent labels identifying pathogen serovars) or biosensors, a gold screen printed electrodes with immobilized CBDs, used for rapid detection of *L. monocytogenes.*	Obeso *et al.*, 2008 Gaeng *et al.*, 2000 Kretzer *et al.*, 2007 de Vries *et al.*, 1999 Tolba *et al*., 2012
surface cleansing	the surface contamination of biofilm forming bacteria on medical devices, food processing machines or pipelines	(1) endolysins CHAPk and LysH5 successfully used for eradication of *S. aureus* biofilms formed on artificial surfaces. (2) LysSMP removed up to 80% of *Streptococcus suis* biofilm, an important animal pathogen and a zoonotic agent.	Fenton *et al.*, 2013 Gutierrez *et al.*, 2014 Meng *et al.*, 2011

To date, endolysins have not yet been approved for human therapy. However, several clinical trials are in process, e.g. *S. aureus* treatment via (1) intra-nasal use of P128 by GangaGen (phase I -safety and phase II- efficacy trial, ClinicalTrials.gov Identifier NCT01746654), (2) intravenous use of PlySs2(CF-301) by ContraFect (phase I trial, ClinicalTrials.gov Identifier NCT02439359), or (3) topical application of Staphefekt™ XDR.300 by Micreos (study phase not provided, ClinicalTrials.gov Identifier NCT02840955). The results have not yet been published.

1.3.2. Polysaccharide depolymerases – novel tools for bacterial biofilms study

To reach biofilm-embedded bacteria for infection and/or destroy the bacterial envelope to facilitate access to the host receptors, phages have to degrade the macromolecular carbohydrates associated with the cell surface (Yan *et al*., 2014; Pires *et al.*, 2016). This is achieved by virion-assoicated polysaccharide depolymerases or by polysaccharide depolymerases released by lysis of other bacteria, that have previously been infected by the phage (Azeredo *et al.*, 2008). These specific enzymes also play important role in the release of bacteriophages trapped within a biofilm (Yan *et al*., 2014). If a phage infects bacteria within a 'young' biofilm and enters the lysogenic state, the maturation of the biofilm may later inhibit the release of phage when it returns to the lytic stage. Indeed, the increased thickness of host biofilms may result in a phage becoming trapped in the basal layers of a biofilm community if it does not possess the enzymatic capacity to liberate itself from the EPS mass that surrounds it (Skillman *et al.*, 1999). To date, this group of enzymes has been reported only in tailed dsDNA phages (Roach *et al.*, 2015).

1.3.2.1. Structure and mode of action

The polysaccharide depolymerases, exhibiting bacteriostatic activity, are highly specific enzymes and can only degrade polysaccharides of a certain composition or structure (Drulis-Kawa *et al*., 2012; Kaur *et al.*, 2015). They can act on the bonds inside (endo) the polysaccharide chain or from the ends (exo) (Barbirz *et al.*, 2008).

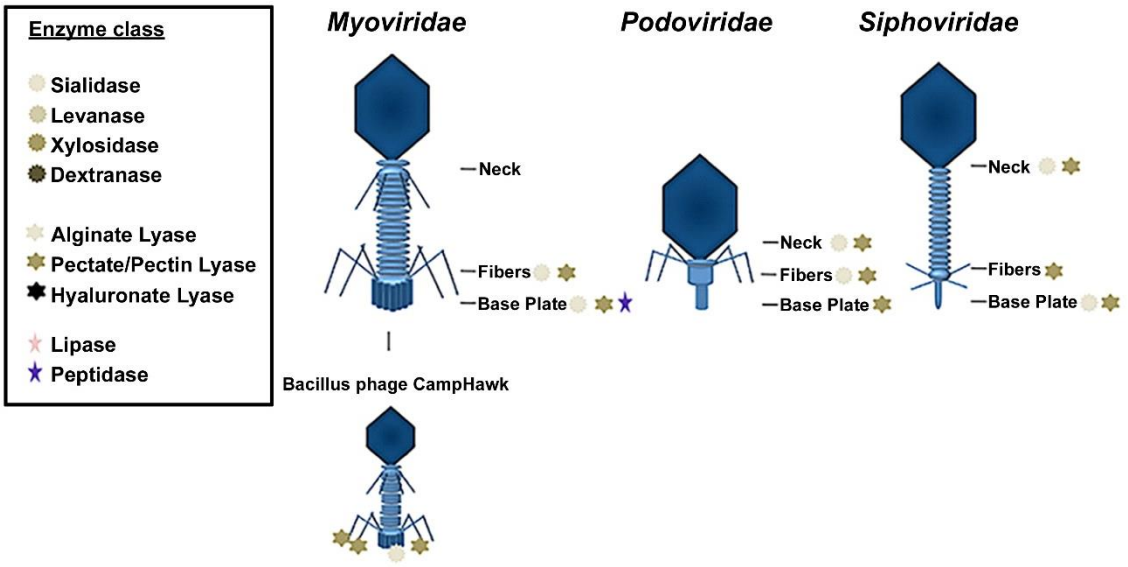


Figure 1.13: Polysaccharide depolymerases distribution through the structure of tailed bacteriophages. Affiliation to different polysaccharide deplymerase classes is presented in the frame. Levanases, xylosidades, hyaluronidases, dextranases, and lipases represent rare depolymerases domains and therefore the location was difficult to identify. The *Bacillus* phage CampHawk represents the only known to date phage, that comprises four depolymerases in different genes (one sialidase and three pectate lyases) (Pires *et al.*, 2016).

During phage path towards the bacterial surface, depolymerases bind to their specific secondary receptors (e.g. capsular polysaccharides) surrounding the phage target cell and degrade the sugar bonds. As soon as the phage reaches the bacterial cell surface, it can attach to its primary receptor and start infection (Fig. 1.14) (Waseh *et al.*, 2010).

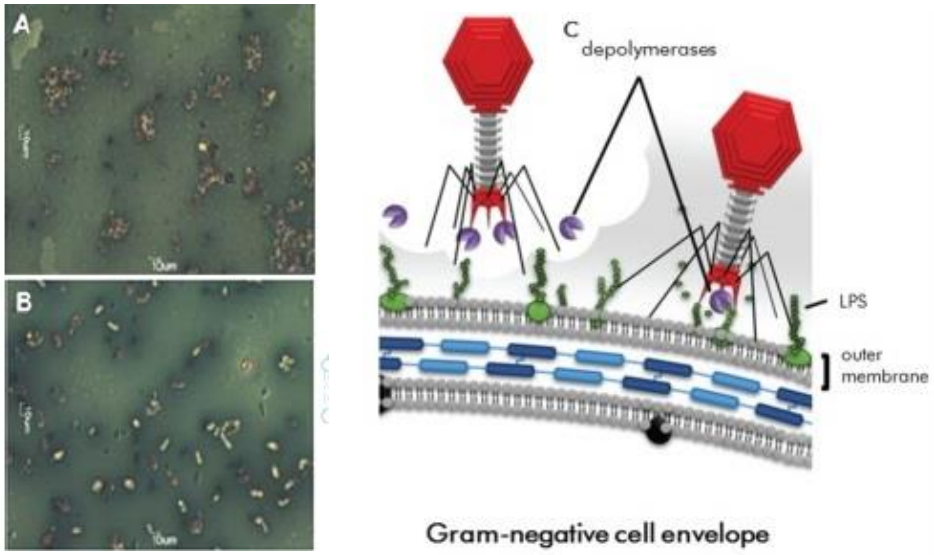


Figure 1.14: The mode of action of polysaccharide depolymerases. Capsule staining of the bacteria in the outside bacterial (A) and in the halo (B) zone after infection caused by lytic *Pseudomonas putida* phage φ15. Within the outside halo zone, bacteria are all surrounded by white capsules and packed closely together. In contrast, bacteria in the halo zone are separated and some of them have lost their white capsules (Cornelissen *et al.*, 2011). (C) The phage depolymerase binds to the macromolecule carbohydrates within extracellular polysaccharides (capsules, biofilm EPS) or lipopolysaccharides (LPS) surrounding bacterial cells (secondary receptors) and degrades them to gain access to the specific primary receptor on the bacterial cell surface. Next, they can inject their genetic material and start the lytic or lysogenic cycle (Hughes *et al.*, 1998; Labrie *et al.*, 2010, Roach *et al.*, 2015).

The polysaccharide depolymerase activity can be observed on a bacterial lawn as a constantly increasing halo zone surrounding the phage plaques (Fig. 1.15) (Cornelissen *et al.*, 2011, 2012; Harper *et al.*, 2014). Bacteria located within haloes areas are decapsulated due to diffusion of EPS-degrading enzymes (Sutherland *et al.*, 2004; Hughes *et al.*, 1998; Lu *et al.*, 2011). However, some phages with active polysaccharide depolymerases do not form clear halo zones, e.g. the *Salmonella* phage PVP-SE1 with pectate-lyase (Santos *et al.*, 2009).

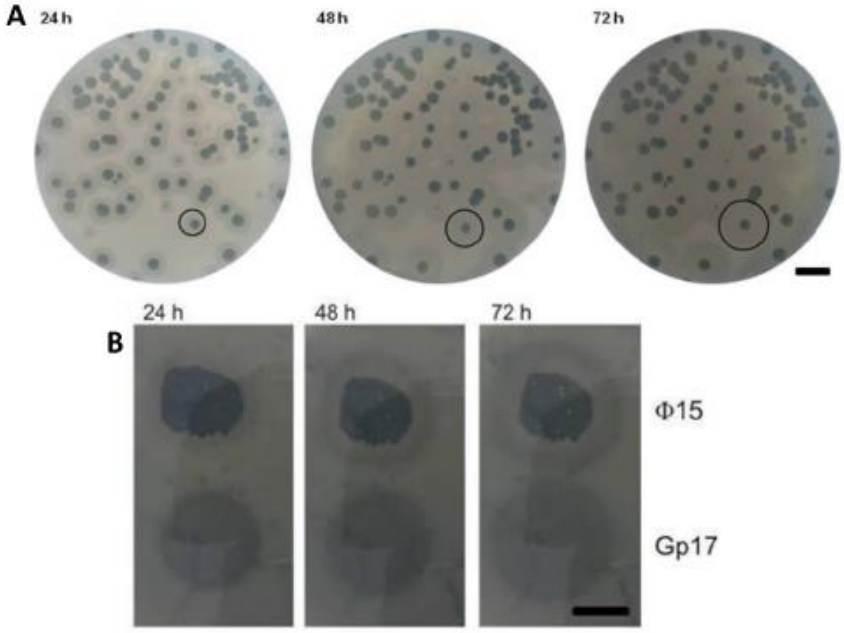


Figure 1.15: Depolymerase activity of phage φ15. (A) Halo formation around phage plaques on a bacterial lawn of *P. putida* PpG1 performed by double layer agar method. (B) Halo formation of phage φ15 (above) and purified tail spike protein (below). The halo created by dropping 10 µl of recombinant purified Gp17 (1 mg/ml) on a bacterial lawn of *P. putida* PpG1 looks similar to the halo around the lysis zone produced by a drop of 10^8 pfu/ml of φ15 phage. The diameter of both halo zones increases over the course of time (A and B). Scale bar represents 1 cm (Cornelissen *et al.*, 2011).

Depending on catalytic activity, phage polysaccharide depolymerases can be classified into two main classes hydrolases (or polysaccharases) and polysaccharide lyases (Sutherland, 2001; Waseh *et al.*, 2010). First hydrolytically break the glycosyl-oxygen bound in the glycosidic linkage. Among this class six different groups can be distinguished: 1) sialidases or neuraminidases, that hydrolyze the α-linkage of the terminal sialic acids of glycans; 2) levanases, that catalyze the hydrolysis of the β-2,6-linked main chain of levan (member of fructans); 3) xylosidases hydrolyzing xylan sugar; 4) dextranases hydrolyzing dextran sugar, 5) rhamnosidases hydrolyzing rhamnogalacturonan sugar, especially interesting are endorhamnosidases that hydrolyze LPS of Gram-negative species and 6) peptidases, that act on peptide bonds, such as poly-γ-glutamate (γ-PGA) hydrolase (Pires *et al.*, 2016; Davies *et al.*, 1995; Kim *et al.*, 2011,

Murakami *et al.*, 1992; Miasnikov, 1997; Sunna *et al.*,1997; Jiménez 2009; Yadav *et al.*, 2010; Kimura *et al.*, 2003). Polysaccharide lyases cleave 1,4- glycosidic bonds between the monosaccharide and the C4 of uronic acid and introduce a double bond between the C4 and C5 of this molecule (Waseh *et al.*, 2010; Drulis-Kawa *et al*., 2012). This class can be divided into three groups: 1) hyaluronidases or hyaluronate lyases, that are able to digest the hyaluronate, a linear unsulfated glycosaminoglucan polymer and cleave the β-1-4 glycosidic linkage by β-elimination; 2) alginate lyases can have mannuronate or guluronate lyases activity, that catalyze the degradation of α-L-guluronate and its C5 epimer β-D-mannuronate (alginate), and 3) pectin/pectate lyases, that degrade galacturonic acid (Pires *et al.*, 2016; Hynes *et al.*, 2000; Wong *et al.*, 2000; Kim *et al.*, 2011; Frirdich *et al.*, 2005; Garron *et al.*, 2010). Furthermore, there is a small pool of phages exhibiting lipases or triacylglycerol hydrolase (degrade the carboxyl ester bonds) activities (Pires *et al.*, 2016; Jaeger *et al.*, 1994; Gupta *et al.*, 2004).

Phage EPS depolymerases can comprise one or two domains in one gene or be separated in different genes, as it is in case e.g. *E. coli* phage ϕK1-5 with two different enzymatic tail fiber proteins (endosialidase and lyase activity) (Scholl *et al.*, 2001).

1.3.2.2. Current state of art

An innovative polysaccharide depolymerase approach exploits the weapons of the bacteriophage. Available literature is limited, focusing mainly on Gram-negative bacteria depolymerases and enzymes of bacterial origin.

Phages intimately coevolved with their host bacteria for 3 billion years, and therefore evolved highly efficient but specific enzymes. They offer a novel stand-alone strategy to control biofilm growth, as well as complementary to existing strategies. Unfortunately, phages exhibiting enzymes with strong polysaccharide depolymerase activity are less abundant, however already few examples can be found in literature. In Tait *et al.* study (2002) the cocktail application of three *Enterobacter cloacae* phages completely eradicated the bacterial biofilm. Hanlon *et al.* (2001) as well as Glonti *et al.* (2010) demonstrated alginate degradation caused by phage polysaccharide depolymerases among biofilms formed by mucoid *P. aeruginosa* strains, facilitating dispersion of the biofilm matrix and its viscosity reduction. Alginate lyase of *P. aeruginosa* phage PT-6 was able to reduce mucus viscosity by 62 - 66% during 15 min treatment. This is an important observation, that may be of interest

for the treatment of CF patients' infection. Furthermore, Gutiérrez *et al.* (2015) research focusing on EPS depolymerase (Dpo7) isolated from a *S. epidermidis* phage for *S. aureus* and *S. epidermidis* biofilms supported the specificity of these enzymes. Dpo7 was able to remove staphylococcal biofilms formed by polysaccharide producer strains, while it failed in polysaccharide-independent biofilms. The oral administration of tail spike protein, P22sTsp specific to *Salmonella*, significantly reduced *Salmonella* infection in chickens. The P22sTsp possess endorhamnosidase activity and/or binding activity of O-antigen in LPS (Waseh *et al.*, 2010; Roah *et al.*, 2015).

On the other hand, also few enzymes of bacterial origin have been examined, such as *B. anthracis* capsule depolymerase (CapD) with a gamma-glutamyl transpeptidase activity. The treatment of *B. anthracis* cells with purified enzyme, caused capsule degradation and increased susceptibility to innate immune responses (e.g. phagocytosis) (Scorpio *et al*., 2010). Another approach is a synergistic action with antibiotics. Bansal *et al*. (2014) successfully treated *K. pneumoniae* mouse infection model with combinatory therapy of bacterial EPS depolymerases and gentamicin, where proteins efficiently removed the bacterial capsule and improved the gentamicin activity. Importantly, bacterial depolymerases have weaker activity compared to their phages counterparts, because they regulate biofilm formation, while phage depolymerases are tailed by nature to degrade bacterial biofilms.

Other approaches include phage engineering, where phages deprived of polysaccharide depolymerases can be armoured with phage enzymes. This was presented in Lu *et al.* (2007) study, where T7 phage was engineered with the dispersin B and gained antibiofilm activity. Another possibility is application of side-directed mutagenesis strategies, such as domain shuffling, which could have possibility to 'relax' their natural specificity and to create broad-spectrum polysaccharide depolymerases.

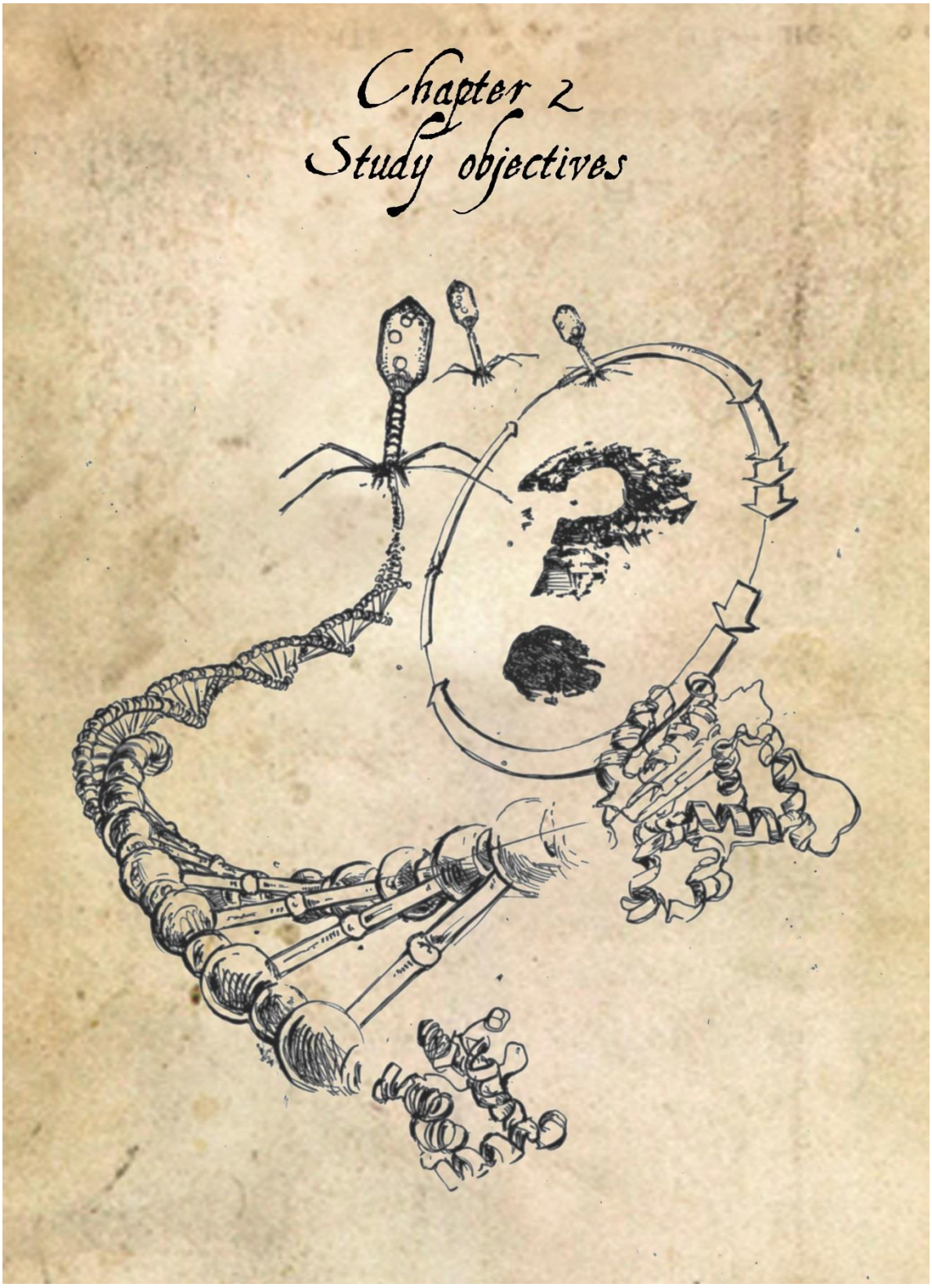
1.3.2.3. Potential applications of phage polysaccharide degrading enzymes

The positive effects of polysaccharide depolymerase administration are promising, however there are still some obstacles, such as: 1) the risk for elimination by the host immune system; 2) lack of stability in the human body; 3) high specificity hindering the multispecies biofilms treatment (Fernebro, 2011). Additionally, the major drawbacks is also low homology in current databases and a small pool of genes with correctly assigned function as well as lack of structural analysis of these enzymes (Drulis-

Kawa *et al*., 2012). From the production perspective, the generation of cost-efficient recombinant protein can also be a challenge.

Nevertheless, these interesting enzymes can have potential applications in medicine, as well as in industry. Xylanases could be used in biobleaching of wood pulp in textile industry (Juturu *et al*., 2012). Rhamnosidases could remove the bitterness from citrus fruit juices or serve as aroma enhancement in wine-making industry (Manzanares *et al.*, 2007). Further, pectin/pectate lyases could assist in fruit juice extraction and clarification (Sieiro *et al.*, 2012). In medicine, alginate lyases have potential to treat cystic fibrosis patient infections in combination with other antimicrobial therapies (Kim *et al.*, 2011; Born *et al.*, 2014). They can facilitate the diffusion of antibiotics, such as aminoglycosides (Hay *et al.*, 2009) through biofilm. Moreover, polysaccharide degrading enzymes can serve as a diagnostic agents, e.g. in neurological tumors, such as neuroblastoma and meningitis, by creation of hybrid proteins consisting the green fluorescent protein (GFP) tag with a catalytically inactive endosialidase, that bind but not degrade polySia (Jokilammi *et al*., 2004; Yan *et al.*, 2014).

Chapter 2
Study objectives



The increasing frequency of multi-drug-resistant strains is particularly concerning as treatment options are severely limited by the absence of effective antibacterials as well as malpractice in the drugs delivery in clinical settings (Breidentein *et al.*, 2011; Poole, 2011). Bacterial infections become severe causing further complications and prolonged hospital treatment, which leads as a consequence to higher tangible as well as intangible costs (Gould, 2006; Haecker, 2009; Hübner *et al.*, 2012).

P. aeruginosa is one of the leading Gram-negative organisms associated with nosocomial infections. As an opportunistic pathogen it is particularly well known for causing infections in immune deficient individuals and presents resistance to many antibiotics (Cornelis *et al.*, 2008). It is responsible for 16.6% of all infections in ICUs (ECDC surveillance report 2014). Despite extensive research to improve the pharmacokinetic and/or pharmacodynamic properties of different classes of antibiotic, only ten new drugs targeting Gram-negatives were introduced in recent years. Unfortunately, six of them are based on combinations of already used antibiotics, to which resistance mechanisms have been already evolved in clinical strains (Mangili *et al.*, 2005). The above results in an urgent need for the development of alternative therapies.

Currently, bacteriophages are in the spotlight. They intimately co-evolved with their host bacteria for 3 billion years, and therefore developed highly efficient predator strategies. There is a variety of approaches for phage therapy implementation in medicine and industry, including use of individual lytic phages or phages cocktails, as well as phage-borne enzymes, such as PG degrading enzymes or polysaccharide depolymerases. Because of a high interest for phages and their products, there is a need for in-depth research of their biology, genome characteristics and antibacterial potential.

The main objective of presented doctoral dissertation is the characterization of bacteriophages lytic towards *P. aeruginosa* as well as two classes of phage-borne enzymes: PG degrading enzymes targeting bacterial cell walls and polysaccharide depolymerases as potential anti-biofilm agents. Achieving the overall aim requires the definition of intermediate objectives:

The characterization of Pseudomonas phages and their therapeutic potential (Chapters 4 through 6)

1) Characterization of basic phage biology, including morphology, host surface receptor, reproduction cycle, viral particle stability and a host range.

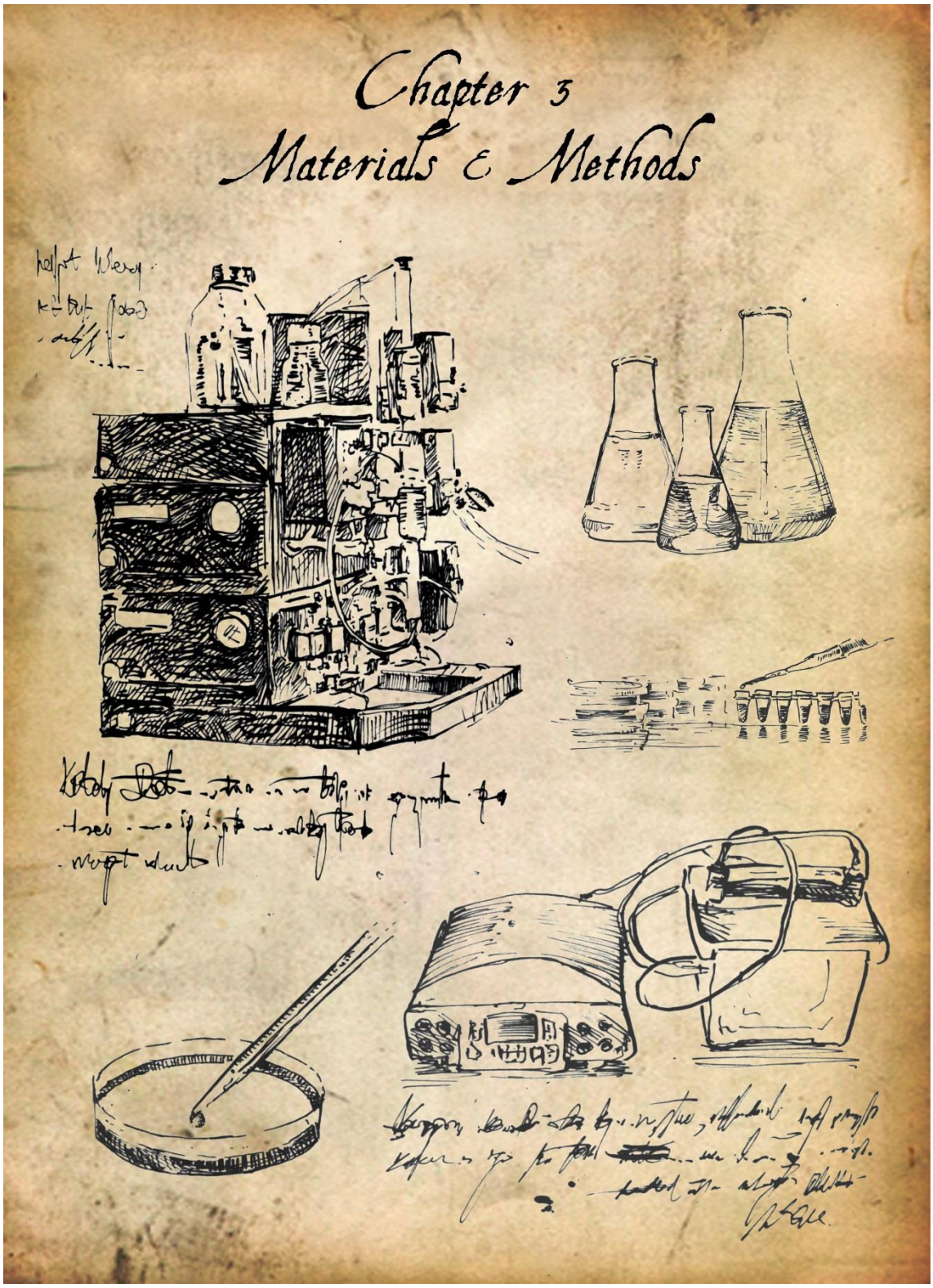
2) Study of phage genome organization and proteome composition and their influence for phage structure and fitness.
3) Comparative genome analysis and construction of protein-sharing networks to reveal phage evolution.
4) Evaluation of phage antibacterial activity *in vitro* with the use of a novel Airway Surface Liquid model on non-CF and CF epithelial cells lines, in an effort to mimic *in vivo* conditions of the respiratory tract.

The exploration of phage-encoded PG degrading enzymes and polysaccharide depolymerases from a functional and application perspective (Chapters 7 and 8)

1) Bioinformatics search of a set of phage genes encoding PG degrading enzymes and polysaccharide depolymerases, and identification of enzymatic activity and structure prediction, based on available homology.
2) Preparation of recombinant proteins, including gene amplification, protein expression optimization and purification to obtain pure and active phage enzymes.
3) Assessment of recombinant protein stability.
4) Evaluation of bactericidal activity and biofilm eradication potential of recombinant enzymes.

The presented project was a part of COST BM1003 Work Group1 activity which focus on phenotypic assessment of bacterial cell surface virulence markers from clinically relevant CF pathogens at different stages of CF lung disease and on the development of new antibacterial therapies.

Chapter 3
Materials & Methods



3.1. Phage manipulations

3.1.1. Microbiology techniques

3.1.1.1. Phage isolation and propagation

The *P.aeruginosa* PAO1 strain was purchased from the American Type Culture Collection (ATCC 15692) and was used as a host for phage propagation. Environmental water samples from irrigated fields in Wrocław, Poland were centrifuged (15,000 × *g* for 15 min) and the supernatant was filtered through a 0.22 μm Millex-GP filter (Merck Millipore) to remove bacterial debris. For phage propagation 0.5 ml of filtered water sample and 5 μl of bacterial host strain in logarithmic phase were added to 5 ml of Mueller Hinton broth (MHB) (Bio-Rad Laboratories) and incubated at 37°C with agitation for 18 h. Every 3, 6, 18 hours each tube was visually inspected. When bacterial lysis occurred, incubation was stopped and sample was centrifuged (5,000 × *g* for 15 min at 4°C). Next, the supernatant was used for large scale phage propagation. An overnight culture of bacterial host strain (5 ml) was added to 500 ml MHB culture and incubated at 37°C with constant shaking. When the bacterial culture reached an optical density at 600 nm (OD_{600}) of ± 0.5, the appropriate phage supernatant was added to infect the host cells. The infected culture was incubated overnight at the same conditions until complete lysis occurred. Next, chloroform was added (10% v/v final concentration, Merck) for an additional 30 min at 37°C and culture was kept overnight at 4°C without shaking. The bacterial debris was further removed by centrifugation (5,000 × *g*, 15 min, 4°C) and supernatant was filtered through a 0.22 μm Millex-GP filter (Danis-Włodarczyk *et al.*, 2015).

3.1.1.2. Phage lysate purification

For phage purification and concentration, first, traditional polyethylene glycol (PEG) precipitation was used, according to the Sambrook and Russell (2001). A 25% PEG_{8000} solution [25% w/v polyethylene glycol, with an average molecular weight of 8,000 Da, 1M NaCl (Acros Organics), demineralized water] was added to filter-sterilized phage lysate to a final concentration of 8% PEG_{8000} and stirred overnight at 4°C. The supernatant was discarded by centrifugation (4,600 × *g*, 45 min, 4°C). The resulting pellet was resuspended in 2 ml of phage buffer (10 mM Tris-HCl, 10 mM $MgSO_4$, 150 mM NaCl, pH 7.5, Acros Organics) and stored at 4°C in the dark. Next, phage lysate was purified with CsCl-gradient ultracentrifugation as described by Ceyssens *et al.*

(2008). In short, the phage solution was layered on a CsCl (MP Biomedicals) step gradient (1.33, 1.45, 1.50 and 1.70 g/cm^3), subsequently centrifuged (140,000 × g, 3 h, 4°C) and dialyzed three times for 30 min against 250 volumes of phage buffer using Slide-A-Lyzer Dialysis Cassettes G2 (Thermo Fisher Scientific Inc.). Purified phages samples were stored at 4°C in the dark.

3.1.1.3. Phage titration

The titer of phage solution was assessed using the classic double-agar layer technique according to methods described previously (Adams *et al.*, 1959). A serial dilution was made and added to a soft agar heated to 56°C [5 g/L yeast extract (LabM), 10 g/L NaCl (Acros Organics), 10 g/L tryptone (LabM) supplied with 0.7% agar, LabM]. This solution was gently mixed and plated on LB agar [5 g/L yeast extract (LabM), 10 g/L NaCl (Acros Organics), 10 g/L tryptone (LabM) supplied with 1.5% agar, LabM]. After overnight incubation at 37°C the phage titer was determined as plaque forming units per ml (pfu/ml). All titrations were performed in triplicate and the result was the average pfu/ml.

3.1.1.4. Transmission electron microscopy

A filtered high-titer phage lysate (10^8 - 10^{10} pfu/ml) was centrifuged at 25,000 × *g* for 60 min. The pellet was washed twice in ammonium acetate (0.1 M, pH 7.0, Acros Organics). Phages were deposited on copper grids with carbon-coated Formvar films and stained for 10 s with uranyl acetate (2%, pH 4.5, Acros Organics) or phosphotungstate (2%, pH 7, Acros Organics). Excess liquid was blotted off and phages were examined using a Zeiss EM 900 electron microscope by technician Sylwia Nowak at the Laboratory of Microscopy Techniques, University of Wrocław, Poland and/or Prof. Hans W. Ackermann, Department of Microbiology-Infectiology and Immunology, Medical School, Laval University, Quebec, Canada. The magnification was calibrated using T4 phage tail length (114 nm) as a standard (Drulis-Kawa *et al.*, 2014).

3.1.1.5. Time - lapse microscopy

To follow phage infection at the bacterial single-cell level, a bacterial host culture, grown to logarithmic phase, was mixed with phage solution in multiplicity of infection (MOI) of 10 and spotted on 0.2% LB agar pads, as described previously by Cenens *et al*. (2013). Infection process was recorded at 37°C immediately post infection in real time

for 5 h with a Nikon Eclipse Ti Time-Lapse Microscope (Nikon, Champigny-sur-Marne, FR) using the NIS-Elements AR 3.2 software (Nikon), available at the Laboratory of Food Microbiology (KULeuven; Prof. A. Aertsen) with the help of Dr. William Cenens. Images were acquired using NIS-Elements (Nikon) and resulting pictures were further handled with open source software ImageJ (Downloaded from http://rsbweb.nih.gov/ij/).

3.1.1.6. Burst size experiments

A one-step growth curve was performed according to the method of Pajunen *et al.* (2000), with modifications. An equal volume of bacterial culture at OD_{600} of 0.4 was mixed with phage suspension (10^6 pfu/ml) to obtain a MOI of 0.01. Phages were allowed to adsorb for 8 min at 37°C, after which the mixture was diluted to 10^{-4}. Triplicate samples were taken every 5 min during 1-1.5 h and titrated (Danis-Włodarczyk *et al.*, 2015).

3.1.1.7. Sensitivity of phage particles to heat, chloroform and pH

The sensitivity to heat was determined by incubating phage suspension (10^8 pfu/ml) in phage buffer at various temperatures (40, 50, 60, 70, and 80°C) and for 5, 15, 30, 45 and 60 min. Chloroform sensitivity was determined by the incubation of equal volumes of phage suspension (10^8 pfu/ml) and chloroform for 1 and 24 h at 4°C with intermittent shaking. The pH stability was tested by incubation of 100 µl of phage suspension (10^8 pfu/ml) in 900 µl of universal buffer (10 mM Tris-HCl pH 7, 10 mM $MgSO_4$, 150 mM NaCl) at pH range 2–12 (pH 1 to 13 adjusted with NaOH, or HCl). Phage titer was assessed after 1h incubation at room temperature as previously described (Danis-Włodarczyk *et al.*, 2015).

3.1.1.8. Phage typing

The lytic activity of isolated viruses was examined on 57 clinical *P. aeruginosa* strains from the Military Hospital Neder-Over-Heembeek, Brussels, Belgium (Pirnay *et al.*, 2003) and compared to the activity of different phages groups: *Lit1virus* (LIT1), *Luz7virus* (LUZ7), *LUZ24virus* (LUZ24), *Kmvvirus* (LUZ19, LKD16, LKA1, φKMV), *Pbunavirus* (LBL3, LMA2, LSL4, KT28, KTN6), *Phikzvirus* (KTN4, φKZ) and a new giant phage PA5oct. Phage suspension in phage buffer (10^6 pfu/ml) was spotted (10µl) on bacterial lawn formed on LB agar. The plates were checked for confluent lysis

or plaque formation against negative, uninfected controls after 4–6 h and again after 18 h. Experiment performed in triplicates (Danis-Włodarczyk *et al.*, 2015, 2016).

3.1.1.9. Phage receptor analysis

The phage specificity to particular bacterial receptors was tested on PAO1 mutants deficient in biosynthesis of A-band and B-band O antigen, flagella, type IV pili, or alginate production (Table 3.1). To determine bacterial susceptibility to phage-mediated lysis, a 10 µl drop of the phage suspension (10^6 pfu/ml) was put on a bacterial lawn and incubated at 37 °C, in three independent experiments. The plates were checked for confluent lysis or plaque formation against negative, uninfected control after 4–6 h and again after 18 h (Danis-Włodarczyk *et al.*, 2015, 2016).

Table. 3.1: List of *P. aeruginosa* PAO1 mutants used for phage receptor identification.

Bacterial strain	Phenotype	Origin
PAO1 (ATCC 15692)	Wild type	American Type Culture Collection
PAO1 Pirnay	Wild type with inactive type IV pili	Military Hospital Neder-Over-Heembeek, Brussels, Belgium
PAO1 Krylov	Wild type	
PAO1 Δrmd (A-, B+)	Deficiency in D-rhamnose biosynthesis; lack of A-band LPS	Department of Molecular and Cellular Biology, University of Guelph, Canada
PAO1 ΔrmLC (A-, B-, core-)	Deficiency in L-rhamnose biosynthesis; truncate core region, lack of A-band and B-band LPS	
PAO1 ΔwaaL (A-, B-)	Lack of WaaL ligating O-polymer to core-lipid A; LPS is devoid of A-band and B-band, semirough (SR-LPS, or core-plus-one O-antigen)	
PAO1 ΔwbpL (A-, B-)	Lack of glucosyltransferase WbpL essential for initiation of both A-band and B-band synthesis	
PAO1 ΔfliC ΔalgC ΔpilA	Lack of flagella; lack of AlgC required for A-band, core oligosaccharide, and alginate biosynthesis; lack of type IV pili	Technical University Hamburg, Germany
PAO1 ΔfliC wt algC ΔpilA	Lack of flagella; lack of type IV pili	
PAO1 ΔfliC wt algC wt pilA	Lack of flagella	
PAO1 wt fliC wt algC wt pilA	Wild type	

3.1.1.10. Phage inactivation by LPS and LPS- binding assay

Experiments were performed by Dr. D. Augustyniak in Department of Pathogen Biology and Immunology (University of Wrocław, Poland). The LPS was isolated with the use of hot phenol/water method by Westphal *et al.* (1965) with modifications. The phage inactivation by LPS was performed as previously described by Kropinski (2009), with modifications. LPS-binding assay was developed by Dr. D. Augustyniak. The detailed description can be found in Danis-Wlodarczyk *et al.* (2015).

3.1.2. Functional analysis

3.1.2.1. Biofilm eradication analysis

The *P. aeruginosa* PAO1 biofilms were formed on µ-Slide 8 Well plates (ibidi GmbH). Overnight bacterial culture was refreshed and grown in LB broth till OD_{600} of 0.2 (10^7 cfu/ml) and 100 µl/well was introduced. Plates were sealed to avoid water loss due to evaporation and biofilms were allowed to grow for 24 h or 48 h at 37 °C, 5% CO_2 without shaking. After 24 h 100 µl of fresh LB broth was added to biofilms. Finally, to assess biofilm degradation by phage particles, 24 h and 48 h old biofilms were treated with 100 µl of the appropriate CsCl purified phage suspension in phage buffer (10 mM Tris-HCl pH 7, 10 mM $MgSO_4$, 150 mM NaCl) and 20 µl of the LIVE/DEAD *Bac*Light stain (LIVE/DEAD *Bac*Light Bacterial Viability Kit, Thermo Scientific Inc.) and incubated for further 24 h at 37 °C. Untreated biofilm served as an negative control and as a positive control cells were treated with 400 µg/ml of gentamicin. Biofilms were visualized by confocal laser scanning microscopy (Zeiss LSM 510 Meta 40× objective, Jena). For quantitative and qualitative analysis of biofilm degradation further experiments were introduced, i.e. interferometry and goniometry performed by Prof. M. Arabski (Department of Microbiology, Institute of Biology, The Jan Kochanowski University in Kielce, Kielce, Poland) and profilometry executed by Nikos Ekizoglou (Zeta Instruments Co., San Jose, USA), as described by Danis-Włodarczyk *et al.* (2015, 2016).

3.1.2.2. Gentamicin exclusion assay on Airway Surface Liquid infection model

NuLi-1 cells (Normal Lung, University of Iowa), derived from human airway epithelium of normal genotype, and a CF cell line, called CuFi-1 (Cystic Fibrosis, University of Iowa), derived from bronchial epithelium of a homozygous CFTR F508del/F508del individual, were kindly provided by Dr. J. Zabner (University of Iowa,

Iowa City, IA). The ASL model was prepared according to methods described by Zabner *et al.* (2003). The experiment was conducted as described by Danis-Włodarczyk *et al.*, 2016. Both cell lines were inoculated with 25 µl of *P. aeruginosa* PAO1 reference strain (6.2×10^7 cfu/ml), nonCF0038 isolate from burn wound (6.5×10^7 cfu/ml) and CF708 small colony variant (1.0×10^6 cfu/ml) at $OD_{600\,nm}$ of 0.1, and incubated for 1.5 h at 37 °C, 5% CO_2. A 25 µl volume of the appropriate CsCl purified phage suspension in phage buffer was added to each millicell hanging cell culture insert. Subsequently, cells were incubated for 1.5 h at 37 °C, 5% CO_2. Next, the cells were washed with PBS and apical washes were serially diluted in DMEM:F12 medium (Sigma-Aldrich) and quantified by viable counts on LB agar (Sigma-Aldrich) after 24 hours. To evaluate the ability of *P. aeruginosa* to invade into epithelial cells and the ability of phage to prevent this invasion, extracellular and adherent bacteria were killed by addition of 400 µg/ml gentamicin (Thermo-Fisher Scientific) and incubated for 1 h at 37 °C, 5% CO_2. The drug was subsequently removed by PBS washing and epithelial cells were lysed with 100 µl of 0.4% Triton-X100 (Sigma-Aldrich) for 15 min at 37 °C, 5% CO_2. The resulting lysate was serially diluted in DMEM:F12 medium and quantified by viable counts on LB agar after 24 hours. Several controls of epithelial cells viability were prepared: (i) a negative control without treatment, (ii) a negative control with TC media; (iii) a positive control containing Triton-X100; (iv) 1.5 h after phage treatment; (v) 1.5 h after strains treatment. The Nuli-1 and CuFi-1 cells were stained with 8 µM Calcien AM (live staining) (Life Technologies) and 3 µM propidium iodide (PI) (dead staining) (Life Technologies) according to manufacturer's instructions. Negative controls (untreated) and positive controls (0.25% Triton-X100 treated) were included in the experimental set up. After staining, in all cases, filter inserts were XZ scanned using a confocal microscope (Zeiss LSM 510 Meta 40× objective, Jena). No toxicity influence of cell lines was noticed for phage and bacterial samples. The data were analyzed using the Statistica software package (StatSoft, Tulsa, OK, USA). All the values were expressed as mean ± SD and significant differences between variations (denoted p-values < 0.05) were found by means of the Snedecor-Fisher test using one-way ANOVA.

3.1.3. DNA manipulation techniques

3.1.3.1. Phage DNA isolation and sequencing

Phage DNA was isolates according to the modified protocol for λ DNA isolation, as previously described by Ceyssens *et al.* (2009). Purified phage particles (10^{10} pfu/ml) were disrupted by 1 h incubation at 56 °C in 0.5% (w/v) SDS (Janssen Chimica), 20 mM EDTA (Acros Organics) and 5 μg/ml proteinase K (Thermo Fisher Scientific Inc.). After a phenol/chloroform DNA extraction and ethanol precipitation, an additional RNase A treatment (100 μg/ml; Roche Applied Science) was used to remove residual RNA. Next, the high throughput sequencing was performed using the Illumina MiSeq platform available at the Nucleomics Core (VIB, Belgium). A 2*150 bp paired-end library (Nextera XT sample prep) was prepared and sequenced. The reads were assembled in a single contig with a 100–6000 fold coverage using CLC genomics Workbench *de novo* assembly algorithm (CLC bio, Cambridge, MA), as described previously (Danis-Wlodarczyk *et al.*, 2015, 2016).

3.1.3.2. Phage RNA isolation and sequencing

P. aeruginosa PAO1 Krylov strain was grown overnight in 5 ml LB medium at 37 °C. Next, cells were diluted 1:100 in 50 ml fresh medium and further grown at 37 °C till OD_{600} reached 0.3 (~1.2 x 10^8 cfu/ml, early exponential phase). This culture was infected further by PA5oct phage at a MOI of 50, that allowed less than 5% of bacterial survivors to remain within five minutes post infection, producing a synchronous infection. Three time points were selected to represent early, middle and late transcription (5, 15 and 25 min). Over the course of infection, 13.5 ml of infected culture (~2.5 x 10^7 cfu/ml) was halted by rapid cooling with a 1.5 ml of stop solution (RNA buffered phenol diluted with 96% ethanol in 1:9 ratio), to inhibit RNA transcription and degradation at the indicated time points. Collected cells were harvested by centrifugation (5,000 × *g* for 20 min) and stored at -80 °C until use. The efficiency of infection was always checked by CFU counts of the infected culture 5 min post infection, which contained 5% surviving cells compared to the non-infected culture. Total RNA was extracted from the resulting cell pellet by resuspending it in 500 μl of TRIzol (Ambion), to inactivate the RNAses produced by the host, and a classical phenol/chloroform extraction was performed followed by ethanol precipitation. Residual genomic DNA was then removed using TURBO DNase® (Thermo Scientific) and rRNA

was depleted with the Ribo-Zero rRNA Removal Kit (Gram-Negative Bacteria) (Illumina Inc.). The remaining RNA was then processed into cDNA libraries using the TruSeq stranded mRNA sample preparation kit and run on an Illumina NextSeq 500 High 75 at the Nucleomics Core (VIB, Belgium). Each sequence library contained more than 10 million reads aligning to non-ribosomal regions of both the phage and host genome. After trimming, sequencing reads were aligned separately to both the phage and host genomes with the CLC Genomics workbench v7.5.1.

3.1.4. *In silico* analysis

3.1.4.1. Genome analysis

The genome analysis was performed as described by Danis-Włodarczyk *et al.* (2015 and 2016). Potential ORFs were identified using GeneMark S (Besemer *et al.*, 2001), GeneMark.hmm (Lukashin *et al.*, 1998), OrfFinder (Sayers *et al.*, 2011) and were manually analyzed. Translated ORFs were compared to known proteins using the BLASTP (Altschul *et al.*, 1990), HHpred (Soding *et al.*, 2005) and HMMER (Finn *et al.*, 2011), providing further insight into the predicted function of proteins. Conserved protein domains were identified using Pfam (Finn *et al.*, 2006), InterPro (Mitchell *et al.*, 2015) and PHYRE2 (Kelley *et al.*, 2015). Putative tRNA genes were searched for using tRNAscan-SE (Lowe *et al.*, 1997). The intergenic regions were screened for the presence of regulatory elements using fuzznuc (Rice *et al.*, 2000) and MEME/MAST (Bailey *et al.*, 1998) and manually inspected to search for putative -35 (TTGACA) and -10 (TATAAT) boxes. Putative factor-independent terminators were identified with ARNold (Naville *et al.*, 2014). Whole analysis performed on DNA-seq was compared to results from RNA-seq, which were conduced according to Ceyssens and co-workers (Ceyssens *et al.*, 2014). Shortly, the total gene reads mapped to the host (*P. aeruginosa* PAO1) and phage PA5oct were used to annotated gene features using the CLC Genomics Workbench 6.0.2. software for comparative analysis. A genome map representation was constructed with CGview (Stothard *et al.*, 2005).

3.1.4.2. Protein-protein sharing network

To represent the genetic relationships of chosen phages with other bacterial and archaeal viruses from NCBI Refseq as a gene (protein)-sharing network was generated by Dr. Ho Bin Jang, in which each predicted protein was clustered into protein families using the ACLAME database (version 0.4) with the database of “viruses” and

an E-value < 0.001 (Leplae *et al.*, 2010; Lima-Mendez *et al.*, 2008). Additionally, for the phages that share significant gene contents, but are absent in the ACLAME database, protein sequences were retrieved from φKZ (NC_004629), φPA3 (HQ630627), 201φ2-1 (NC_010821), EL (NC_007623), OBP (NC_016571), φJM-2012 (JQ340088), SPN3US (NC_027402), CR5 (NC_021531), φEaH2 (NC_019929), JG024 (NC_017674), PB1 (NC_011810), 14–1 (NC_011703), NH-4 (NC_019451), SPM-1 (NC_023596), LMA2 (NC_011166), LBL3 (NC_011165), KPP12 (NC_019935), SN (NC_011756), and ECML-117 (JX128258). The proteins that could not be assigned into any ACLAME protein families were defined as the unclassified protein families (UPFs) as previously described (Jang *et al.*, 2013). We accepted the transitive nature of sequence families (Casjens, 2003), i.e., a sequence is added to a cluster if it shares a reciprocal best hit relationship with at least one of the sequences of the cluster. The degree of similarity between other phages was generated as the minus logarithmic score by multiplying hypergeometric similarity *P*-value by the total number of pairwise comparisons (Lima-Mendez *et al.*, 2008).

For the giant phage PA5oct analysis, protein sequences were subjected to all-to-all BLASTP searches, with an E-value threshold < 0.001, and defined as the homologous protein clusters (PCs) in the same manner as previously described (Lima-Mendez *et al.*, 2008; Roux *et al.*, 2015). Based on the number of shared PCs between the genomes, the degree of similarity was calculated using vContact (https://bitbucket.org/MAVERICLab/vcontact), as the negative logarithmic score by multiplying hypergeometric similarity P-value by the total number of pairwise comparisons. The protein-sharing relationships among the genomes with similarity values ≥ 2 (E-value ≥ 0.01) were represented (Lima-Mendez *et al.*, 2008). Afterwards, the network was visualized with Cytoscape (version 3.1.1; http://cytoscape.org/), using an edge-weighted spring embedded model, which places the genomes or fragments sharing more PCs closer to each other. In addition, the Markov clustering algorithm (MCL) was used to group each genome into viral cluster (VC). To maximize the cluster homogeneity, the optimal inflation factor was determined by exploring values ranging from 1.0 to 5 by steps of 0.2. The taxonomic affiliation was taken from the ICTV (http://www.ictvonline.org/virusTaxonomy) and/or NCBI taxonomy (http://www.ncbi.nlm.nih.gov/taxonomy). Topological properties of the network were estimated with the Network Analyzer 2.7 Cytoscape plug-in (Brohee *et al.*, 2008).

3.1.5. ESI-MS/MS phage proteome analysis

Phage proteins were extracted from a purified phage suspension by a single methanol/chloroform extraction (1:1:0.75, v/v/v) (Acros Organics) and subsequently precipitated by addition of an equal volume of methanol (16,060 × g, 6 min). The dried phage protein pellet was resuspended in 4x SDS-PAGE loading buffer [200 mM Tris pH 6.8, 8 mM EDTA, 40% (v/v) glycerol, 4% (w/v) SDS, 0.4 % (w/v) bromophenol blue] and boiled for 5 min before loading onto a 12% SDS-PAGE gel. Protein gels were stained afterwards with GelCode Blue Safe (Thermo Scientific). Further, the entire lane of a phage protein profile was prepared for ESI-MS/MS as previously described by Van den Bossche *et al.* (2014) and Ceyssens *et al.* (2014). Shortly, the SDS-PAGE gel was cut into slices, that were subjected to trypsin digestion (Shevchenko *et al.*, 1996) prior to mass spectrometry analysis. Protein digests were analyzed by ESI-MS/MS on a LCQ Classic (ThermoFinnigan) equipped with a nano-LC column switching system (Prof. J-P Noben, UHasselt) as described in Dumont *et al.*, 2004. The analysis of the mass spectrometric RAW data was carried out using Proteome Discoverer software v.1.3 (Thermo Scientific) with build-in Sequest and interfaced with an in-house Mascot v.2.4 server (Matrix Science).

3.2. Protein characterization techniques

3.2.1. *In silico* analysis

The protein domains were identified by BLASTP, HHPRED, HMMER, PHYRE2 and InterPro. The PDBsum database was used for protein secondary and tertiary structure predictions (de Beer *et al.*, 2014), that was further manually analyzed with the use of the PyMole and the MOE programs. Clustal W2 and Clustal Omega were used for protein sequence comparison (Larkin *et al.*, 2007; Sievers *et al.*, 2011). The ligand binding site could be predicted with the use of the 3DLigandSite-Ligand binding site prediction server (Wass *et al.*, 2010).

3.2.2. DNA manipulations and cloning

3.2.2.1. Polymerase chain reaction

Purified genomic DNA of *Pseudomonas* phages, KT28, KTN6, PA5oct, KTN4, LUZ7 and LKA1 served as a template for amplification of the different ORFs in a polymerase chain reaction (PCR). Reactions were performed with 0.02 units/µl

of Phusion High-Fidelity DNA Polymerase (Thermo Scientific Inc.), 200 μM of each dNTP (Thermo Scientific Inc.) and 0.5 μM of both primers (Table 3.2) in the supplied 5X Phusion HF buffer and subjected to 35 cycles of amplification using a Trio Thermoblock (Biometra). Primer annealing temperatures were adapted to the specific primer pair and extension times on the length of the PCR-product. Next, single deoxyadenosine (A) was added to PCR products to the 3′ ends by 0.5 units of DreamTaq™ polymerase (Thermo Scientific) to enable TOPO® T/A cloning. The PCR products were separated by 1% agarose gel electrophoresis in TAE electrophoresis buffer (40 mM Tris-HCl pH 7.2, 500 mM sodium acetate, 50 mM EDTA, Acros Organics). DNA fragments were stained with ethidium bromide and visualized by UV illumination. To estimate the size and the quantity, a reference λ/*Pst*I ladder was run in parallel with the fragments.

Table 3.2: The overview of constructs and corresponding cloning conditions. The first column refers to the name of the encoded protein. For each construct, the vector together with cloning site and primers for PCR reaction are listed. The PCR reaction was always performed on phage genomic DNA. The ATG start triplet and TAA/TGA/TAG stop triplets are written in italic for the *forward* and *reverse* primers with stop codon, respectively.

Insert	Vector and Cloning site	Primers
LKA1 gp49 *	pEXP5-NT/TOPO ®, * pBAD/Thio TOPO® TA-cloning	*ATG*GCGCAAACACCCAG* *TCA*CACCTCATAAATACCTG* CACCTCATAAATACCTG*
LKA1 gp49 domain	pEXP5-NT/TOPO®, pEXP5-CT/TOPO® pBAD/Thio TOPO® TA-cloning	*ATG*GGACTGTTAGTGGAGGCTG *TTA*CGTTGTATTTGTAGTGAGGAC CGTTGTATTTGTAGTGAGGAC
LUZ7 gp56	pEXP5-NT/TOPO®, pEXP5-CT/TOPO® pBAD/Thio TOPO® TA-cloning	*ATG*ACCACAAAGGTGATCTTC *TTA*GTACCACCTCCCGATGGCCAT GTACCACCTCCCGATGGCC
LUZ7 gp56 domain	pEXP5-NT/TOPO®, pEXP5-CT/TOPO® pBAD/Thio TOPO® TA-cloning	*ATG*AATGGCACCTGGCAGAAGG *TTA*ACCCGTGAACGACGGACTC ACCCGTGAACGACGGACTCT
KT28 gp41	pEXP5-NT/TOPO®, pEXP5-CT/TOPO® pBAD/Thio TOPO® TA-cloning	*ATG*TCTAGTGATTTGGATGAATT *TCA*GTAGAGTATCTTAGTTGCG GTAGAGTATCTTAGTTGCGG
KT28 gp41 domain	pEXP5-NT/TOPO®, pEXP5-CT/TOPO® pBAD/Thio TOPO® TA-cloning	*ATG*GGCGCCCAGTCTCTT *TAA*TTCCATCAGCTCGATTG TTCCATCAGCTCGATTGGAG
KT28 gp29	pEXP5-CT/TOPO® TA-cloning	*ATG*AAAATCACGAAGGATGT GGAGCCTTGATTGATCGC
KTN6 46	pEXP5-CT/TOPO® TA-cloning	*ATG*AAAATCACAAAGGATATTCTTATC GGAGCCTTGATTGATCGC
PA5oct gp214	pEXP5-CT/TOPO® TA-cloning	*ATG*AGTACATATTCTAGAACAGTAGGTG TGCTCCTTTATTTTTGTGCTC
PA5oct gp250	pEXP5-NT/TOPO® TA-cloning	*ATG*GATAATAATATTAGAAAAAAATTAATTGC ATTTAAGTTTCCTAAACATACATTCC
KTN4 gp48	pEXP5-CT/TOPO® TA-cloning	*ATG*ACTGAATCAACTACTTTTAGTTTTACC TTTTACTTTAGCAGTTGATACTTCACATAC

(*) Full length LKA1 gp49 construct made and described by A. Corenlissen, 2011.

3.2.2.2. Plasmids and cloning

The commercially available pEXP5-NT/TOPO®, pEXP5-CT/TOPO® and pBAD/Thio TOPO® vectors (Invitrogen) are provided as linearized vectors, which allows a direct TA-cloning strategy due to the covalently coupled topoisomerase I (Fig. 3.1, 3.2). In these vectors, genes are fused to a N- or C-terminal 6x His-tag respectively, which allows the expressed proteins to be purified on a Ni^{2+}-NTA (nitrilotriacetic acid) column by affinity chromatography. In case of pBAD/Thio TOPO® vector recombinant proteins are expressed also as fusions to N-terminal thioredoxin, introduced for high-level expression and simple purification. The TA-cloning into a TOPO® vectors was conducted according to the manufacturer protocols. For ORFs cloned into the pEXP5-CT/TOPO® and pBAD/Thio TOPO® expression vectors no stop codon was included in the reverse primer, whereas those intended for pEXP5-NT/TOPO® had a stop codon in the reverse primer. The PCR products were checked for correct length by gel electrophoresis with 1% (w/v) agarose [0.70 g agarose (Sigma-Aldrich), 70 mL TAE-buffer (40 mM Trizma base (Acros Organics), 20 mM acetate (Merck), 1 mM EDTA (Sigma-Aldrich)]. All constructs selected for further analysis and their corresponding details concerning cloning are summarized in Table 3.2. In case of pEXP5-NT/TOPO®, pEXP5-CT/TOPO® vectors transcription of the insert is strictly controlled by a strong bacteriophage T7 promoter. In contrast, pBAD/Thio TOPO® vector uses the araBAD promoter (pBAD).

3.2.2.3. Transformation into TOP10 *E. coli* competent cells

For constructs storage and clone analysis *E. coli* TOP10 cells [genotype F– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ (*araleu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G] were used since they allow stable replication of high-copy plasmids. Chemically competent *E. coli* TOP10 cells were prepared according to the Rubidium Chloride Method (New England Biolabs). The transformation efficiency was verified with the pUC18 vector (Thermo Scientific). Each construct was gently mixed with 50 µl chemically competent *E. coli* TOP10 cells and incubated for 30 min on ice to enable attachment of the DNA to the bacterial membrane. Next, a 30 sec heat shock at 42 °C was applied to enable the introduction of DNA into the cell. Subsequently, 250 µl of SOC medium [Super Optimal broth with Catabolite repression [5 g/L yeast extract (LabM), 5 g/L NaCl (Acros Organics), 20 g/L tryptone (LabM), 0.1 M $MgSO_4$ (VWR International), 0.1 M $MgCl_2$ (VWR International), 4% (w/v) D-glucose (VWR

International)] was added and the cells were incubated for 1 h with shaking at 37°C to enable recovery and expression of the antibiotic resistance genes. Finally, cells were plated out on LB agar containing 100 μg/ml ampicillin (VWR International) as selection pressure for plasmid maintenance and 2% (w/v) D-glucose (VWR International) for catabolite repression of the *lac* operon. Plates were incubated overnight at 37 °C. Single colonies were picked and colony PCR was performed using DreamTaq™ polymerase (Thermo Scientific) and one primer of the gene and one primer of the vector to enable identification of colonies with the correct construct by gel electrophoresis. Glycerol stocks for storage at -20 °C were made by mixing 800 μl overnight culture of the appropriate colonies with 200 μl of 100% glycerol (Acros Organics). Plasmids were isolated using the GeneJET Plasmid Miniprep kit (Thermo Scientific) according to the manufacturer protocol. The sequence of the plasmids was determined by 'LIGHTrun sequencing' (GATC Biotech) and analyzed by Sequencher software version 4.8 (Gene Codes Corporation).

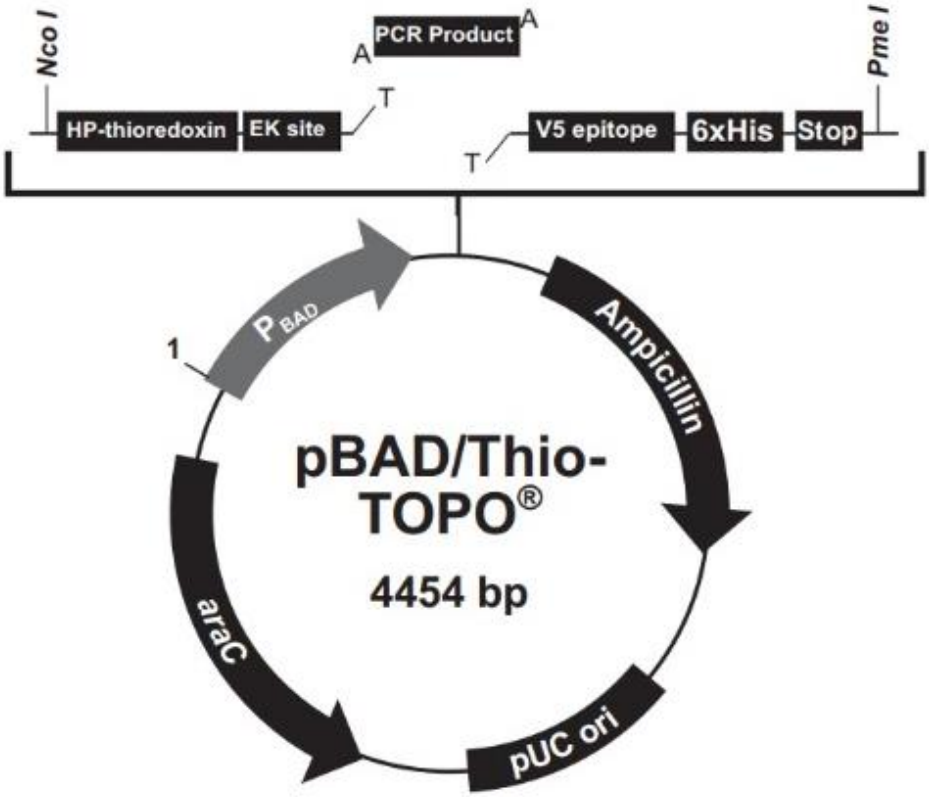


Figure 3.1: Map and features of pBAD/Thio TOPO® vector. The araBAD promoter (P_{BAD}) provides tight, dose-dependent regulation of heterologous gene expression. HP-thioredoxin allows a highly efficient fusion partner for translation of the fusion protein. TOPO® Cloning site enable quick insertion of PCR product for expression. The C-terminal V5 epitope tag (Gly-Lys-Pro-Ile-Pro-Asn-Pro-Leu-Leu-GlyLeu-Asp-Ser-Thr) grants detection of the fusion protein by the Anti-V5 Antibody or the Anti-V5-HRP Antibody. Polyhistidine (6xHis) tag is further required for detection of the recombinant fusion protein with purification with metal-chelating resin (i.e. Ni-NTA). (STOP) rrnB transcription termination region. Ampicillin resistance gene (β-lactamase) is used for selection of the plasmid in *E. coli*. pUC origin provides replication and growth in *E. coli* and araC gene encodes the regulatory protein for tight regulation of the PBAD promoter.

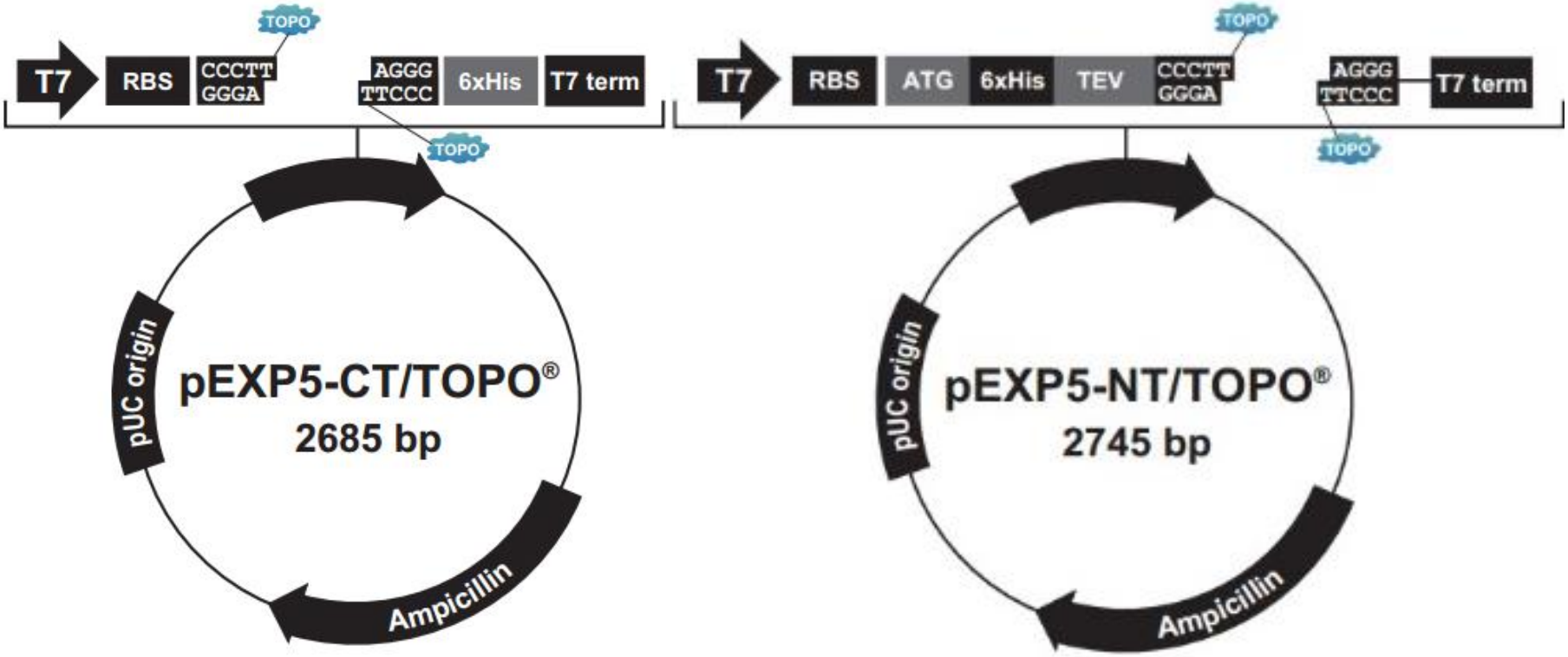


Figure 3.2: Maps and features of pEXP5-NT/TOPO® and pEXP5-CT/TOPO® vectors. From left to right pEXP5-CT/TOPO®, pEXP5-NT/TOPO® maps. T7 promoter provides a high-level expression of recombinant protein in *E. coli* strains expressing the T7 RNA polymerase. T7 forward priming site (T7) allows sequencing in the sense orientation. Ribosome binding site (RBS) is optimally spaced from the initiation ATG for efficient translation of the PCR product. Initiation (ATG) enable translation initiation of the recombinant fusion protein. Polyhistidine (6xHis) tag is further required for detection of the recombinant fusion protein with purification with metal-chelating resin (i.e. Ni-NTA). (TEV) recognition site (Glu-X-X-Tyr-X-Gln-Ser) that allows removal of the N-terminal tag from recombinant fusion protein using TEV protease. TOPO® Cloning site is used for rapid cloning of Taq-amplified PCR product. T7 transcription terminator (T7 term) from bacteriophage T7 that allows efficient transcription termination. Ampicillin resistance gene (β-lactamase) grants selection of the plasmid in *E. coli*. (pUC origin) is ori of replication for high-copy replication and maintenance of the plasmid in *E. coli.*

3.2.3. Recombinant expression

3.2.3.1. Bacterial strains and manipulations

Five different *E. coli* strains [Lemo21(DE3), BL21-Codon Plus(DE3), ArcticExpress(DE3), BL21(DE3)pLysS, BL21 A1] were used to optimize expression of the recombinant proteins (Table 3.3). Therefore, correct plasmid constructs were freshly transformed in these chemically competent *E. coli* cells as described in section 3.2.2.3.

Table 3.3: Characteristics of *E. coli* expression strains and antibiotics requirements for protein expression.

Strain name	Genotype	Features	Features	Additional requirements
Lemo21(DE3)	fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo (CamR) λ DE3 = λ sBamHIo ΔEcoRIB int:lacI::PlacUV5::T7 gene1) i21 Δnin5 pLemo = pACYC184-PrhaBAD-lysY	- lysogenic for λ-DE3 - tunable level of T7 lysozyme by L-rhamnose, that can alleviate inclusion body formation or growth inhibitory effects from toxic proteins	100 µg/mL Amp 30 µg/mL Cm	L-rhamnose (in expression medium
BL21-Codon Plus(DE3)	E. coli B F- ompT hsdS(r-mB) dcm+ Tetr E. coli gal λ (DE3) endA Hte [argU ileY leuW Camr]	- lysogenic for λ-DE3 - T7 expression system - tRNAs for *E. coli* codons that frequently limit translation	100 µg/mL Amp 34 µg/mL Cm	-
ArcticExpress (DE3)	E. coli B F– ompT hsdS(rB-mB) dcm+ Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr]	- lysogenic for λ-DE3 - T7 expression system - cold-adapted chaperonins	100 µg/mL Amp 20 µg/mL Gm	low temperature cultivation (10-13°C)
BL21(DE3) pLysS	E. coli B F- dcm ompT hsdS(r-mB) gal λ(DE3) [pLysS Camr]	- lysogenic for λ-DE3 - T7 lysozyme	100 µg/mL Amp 34 µg/mL Cm	-
BL21 A1	E. coli B F– ompT gal dcm lon hsdSB(rB–mB–) [malB+]K-12(λS) araB::T7RNAP-tetA	- T7 RNA polymerase into *araBAD* operon	100 µg/mL Amp	L-arabinose (at time of induction)

Amp = ampicillin; Cm = chloramphenicol; Gm = gentamicin; λ-DE3 = lysogen construct with T7 RNA polymerase under control of lacUV5 promoter.

Protein expression in BL21(DE3) pLysS competent cells is influenced by the presence of T7 lysozyme, which reduces the background of the targeted protein expression without disturbing the inducer activity of IPTG (isopropyl-1-thio-β-D-galactopyranoside). It suppresses the activity of T7 RNA polymerase, reduces the basal level of protein expression from the gene of interest and increases the tolerance

of the *E. coli* cells against the toxicity. BL21(DE3) pLysS lacks both the Lon and OmpT proteases, which may degrade expressed proteins.

Lemo21(DE3) consist the T7 RNA polymerase, encoded by the lambda DE3 prophage present within the chromosome, expressed from the *lac*UV5 promoter, which is less sensitive to catabolite repression than the wt *lac* promoter. Furthermore, DE3 strains may exhibit uninduced target protein expression, which is remedied by expression of T7 lysozyme (LysY). Membrane protein expression and protein export in *E. coli* are both limited by the throughput capacity of the Sec translocase and in some cases the Tat translocase. T7 expression of proteins targeted to the Sec translocase often leads to accumulation of inclusion bodies or inhibition of cell division, if expression is not regulated. This is why Lemo21(DE3) has a tunable level of T7 lysozyme, controlled by L-rhamnose, which can alleviate inclusion body formation or growth inhibitory effects from toxic proteins. In many cases, decreased expression can result in more protein of interest produced in the desired form. Cells naturally lack the Lon protease, which can degrade recombinant proteins.

The efficiency of heterologous protein high-level expression in *E. coli* is frequently limited by the scarcity of certain tRNAs, that are abundant in the organisms from which the proteins are derived. BL21-CodonPlus(DE3) strain overcomes codon bias as a result of the extra copies of tRNAs genes that frequently limit translation in *E. coli*. Thus a high-level expression of heterologous recombinant genes can be improved in BL21-CodonPlus cells in compare to conventional BL21 strains. Cells lack the Lon and the OmpT proteases.

Further, the ArcticExpress competent cells address the common protein expression hurdle, protein insolubility. They are engineered to increase the yield of soluble protein produced in *E. coli* in low-temperature cultivation, which slows down protein expression. The important role in protein folding play *E. coli* chaperonins, especially the complex GroEL/ES, that retains only about 30% refolding activity at 12 °C, compared to its optimal activity at 30 °C. The ArcticExpress competent cells co-express the cold-adapted chaperonins Cpn10 and Cpn60 from the psychrophilic bacterium, *Oleispira antarctica*, and have 74% and 54% amino acid identity to the *E. coli* GroEL and GroES chaperonins, respectively. They allow for a potent protein refolding at low temperatures of 4–12°C, which potentially may increase the yield of active, soluble recombinant protein. Cells also lack Lon and OmpT proteases.

Finally, BL21-AI™ competent *E. coli* strain offers the maximum protein expression with the tightest regulation available for a T7 expression system. The BL21-AI™ strain also does not contain the Lon and OmpT proteases. In contrast to above strains expression of T7 RNA polymerase is regulated by the araBAD promoter and depends on L-arabinose concentrations.

3.2.3.2. Expression conditions

Freshly transformed *E. coli* expression cells were inoculated in LB broth (1/20 of the expression medium volume) and grown overnight at 37°C with shaking. Next, cultures were inoculated in the expression medium with proper antibiotics. For *E. coli* Lemo21(DE3) cells, also 0.5 mM L-rhamnose (Sigma-Aldrich) was added. Cultures were grown to OD_{600} of ± 0.6 and induced with IPTG or arabinose in case of pBAD/Thio TOPO® vector (GE Healthcare). For *E. coli* BL21 A1, also 0.2% (w/v) L-arabinose (Sigma-Aldrich) was added at time of induction. Next, cells were allowed to grow and express the recombinant protein for 4.5 h at 37 °C or overnight at 16 °C with shaking. For expression in autoinduction medium, cells were grown to OD_{600} of ± 0.8 and further incubated at 16 °C with shaking for exactly 24 h. Media used for expression optimization were as follows: LB [5 g/L yeast extract (LabM), 10 g/L NaCl (Acros Organics), 10 g/L tryptone (LabM)], 2xTY [16 g/L tryptone (LabM), 10 g/L yeast extract (LabM), 5 g/L NaCl (Acros Organics)], modified LB [15 g/L tryptone (LabM), 8 g/L yeast extract (LabM), 5 g/L NaCl (Acros Organics), pH 7.8], TB [12 g/L tryptone (LabM), 24 g/L yeast extract (LabM), 2.3 g/L KH_2PO_4 (VWR Int.), 16.4 g/L K2HPO4 (Acros Organics), 0.4% (v/v) glycerol (Acros Organics)], autoinduction medium [10 g/L tryptone (LabM), 5 g/L yeast extract (LabM), 0.1% (v/v) 1 M $MgSO_4$ (Sigma-Aldrich), 2% (v/v) 50 x 5052-solution (25% (w/v) glycerol (UCB), 2.5% (w/v) D-glucose (Acros Organics), 10% (w/v) D-lactose monohydrate (Sigma Aldrich), 5% (v/v) 20 x NPS-solution (6.6% (w/v) $(NH_4)_2SO_4$ (Acros Organics), 13.6% (w/v) KH_2PO_4 (VWR Int.), 14.2% (w/v) $Na_2HPO_4.2H_2O$ (Acros Organics)].

3.2.3.3. Protein presence analysis

After expression, the cell culture was centrifuged (4000 × *g*, 10 min, 4 °C) and the pellet was stored at -80°C until purification. The presence of recombinant protein was checked by SDS-PAGE and Western Blot. Therefore, 10 ml sample taken post expression was centrifuged (4000 × g, 10 min, 4°C) and the pellet was resuspended in 1 ml lysis

buffer, respectively. The 1 mM Pefabloc® SC (Merck-Millipore) was added to inhibit serine proteases, that could damage the recombinant protein. Next, cells were disrupted by three times freeze-thawing in liquid nitrogen and sonication (6 min, 30 s pulse/30 s rest, 40% amplitude; Vibra-Cell™, Sonics & Materials Inc.) in the presence of 10 µg/ml DNase I (Invitrogen) to reduce the sample viscosity. Total, soluble and insoluble fractions were separated and collected by centrifugation (18,000 × *g*, 30 min, 4 °C). Not induced fraction was harvested from 1 ml sample taken before expression culture induction. Sample was centrifuged (4,000 × *g*, 10 min, 4°C) and the pellet was resuspended in 1 ml lysis buffer. All samples were mixed with 4 x loading buffer and boiled for 5 min at 95 °C prior to loading on the SDS-PAGE gel. The protein concentration was determined by spectrophotometric measurement of absorption at 280 nm using Nanodrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA). The protein molecular weight and extinction coefficient (ε) was determined using EXPASY-ProtParam.

3.2.3.4. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

The purity and composition of the protein samples were visualized by SDS-PAGE. Both 12% and 15% (v/v) polyacrylamide separation gels were prepared, depending on the size of the proteins. Therefore, respectively 3.6 and 3.25 ml MQ water (Merck-Millipore), 2 and 2.5 ml 4 x Tris/SDS buffer pH 8.8 [1.5 mM Trizma base (Acros Organics), 0.4% (w/v) SDS (Janssen Chimica)], 2.4 and 3.75 ml 40% (v/v) acrylamide (Bio-Rad), 0.01% (v/v) ammonium persulfate (APS; Sigma-Aldrich), 0.001% (v/v) N,N,N',N'-tetramethylethylenediamine (TEMED; IBI Scientific) were mixed for the preparation of 2 gels. The two stacking gels were prepared by mixing 2.1 ml MQ water (Merck-Millipore), 800 µl 4x Tris/SDS buffer pH 6.8 [1.5 mM Trizma base (Acros Organics), 0.4% (w/v) SDS (Janssen Chimica)], 300 µl 40% (v/v) acrylamide (Bio-Rad), 0.01% (v/v) APS (Sigma-Aldrich) and 0.001% (v/v) TEMED (IBI Scientific). Gels were placed in the reservoirs filled with SDS-PAGE running buffer [25 mM Trizma base pH 8.3 (Acros Organics), 192 mM glycine (Sigma-Aldrich), 0.1% (w/v) SDS (Janssen Chimica)] and 10 µl of each sample was loaded. Also 7 µl of PageRuler™ Prestained Protein Ladder (Thermo Scientific) was loaded as a reference to estimate the molecular weight (MW) of the proteins. Gels were run at a constant voltage of 200 V and subsequently stained with the GelCode™ Blue Safe Protein Stain (Thermo Scientific).

3.2.3.5. Western blot analysis

Western blotting was performed to verify the presence of the 6 x His-tag in the expressed proteins. The protein fractions (5 µl/well) were loaded on a 12% or 15% (v/v) polyacrylamide gel and SDS-PAGE was performed. Next, proteins were transferred onto an Amersham Hybond P 0.45 PVDF blotting membrane (GE Healthcare) in a presence of a transfer buffer [600 ml MQ (Merck-Millipore), 200 ml 5x Tris-glycine buffer (15 g/L Trizma base (Acros Organics), 75 g/L glycine (Sigma-Aldrich), 200 mL 100% ethanol (Fisher Scientific, Loughborough, UK)]. The membrane was blocked for 1 h at room temperature with PBST buffer [100 ml 10 x PBS, phosphate buffered saline, [80 g/L NaCl (Acros Organics), 2 g/L KCl (Janssen Chimica), 14.4 g/L Na_2HPO_4 (Merck), 2.4 g/L KH_2PO_4 (Merck)], 0.1% (v/v) Tween 20, pH 7.5 (Sigma-Aldrich), 900 ml MQ water (Merck-Millipore)], containing 5% (w/v) powdered milk (Nestlé NAN Pro 2) to avoid unspecific binding of the antibodies and background signals. Afterwards, the membrane was incubated for 1 h with a 1:3000 dilution of mouse monoclonal Anti-polyhistidine antibodies (Promega, Madison) in 10 ml PBST buffer containing 1% (w/v) milk powder at room temperature with gentle shaking. After three washing steps with PBST buffer, the membrane was incubated for 1 h at room temperature with a 1:5000 dilution of Anti-Mouse IgG (H+L) antibodies conjugated to horseradish peroxidase (HRP) (Promega). Subsequently, three washing steps with PBST buffer were performed. Revelation was carried out by enhanced chemiluminescence (ECL) using the Pierce™ ECL Western Blotting Substrate kit (Thermo Scientific). The chemiluminescence signal was trapped on an Amersham Hyperfilm ECL (GE Healthcare).

3.2.4. Protein purification

3.2.4.1. Sample preparation

Expression culture pellets were resuspended in 10 ml lysis buffer [pH 8, 50 mM NaH2PO4 (Merck), 300 mM NaCl (Acros Organics), 10 mM imidazole (Acros Organics)] containing 30% (v/v) glycerol (Acros Organics) to improve stability and 1 mM Pefabloc® SC (Merck-Millipore) was added. The suspension was frozen and thawed three times using liquid nitrogen. After addition of 5 µl 10 µg/mL DNase I (Invitrogen), the suspension was sonicated (6 min, 30 s pulse/30 s rest, 40% amplitude; Vibra-Cell™) to lyse the cells further. Next, the suspension was centrifuged (4600 × g, 50 min, 4°C) and the supernatant was collected.

3.2.4.2. Ni^{2+}-NTA affinity chromatography

The protein lysate, collected from large scale expression, was incubated with 1 ml Ni^{2+}-NTA Superflow (Qiagen) for 1 h rotating at 4°C. After short centrifugation (300 × g, 5 min, 4°C), the resin was gently resuspended in 10 ml lysis buffer with 30% (v/v) glycerol and loaded on chromatographic columns (diameter 25 mm, height 180 mm, CHEMLAND). Subsequently, the resin was washed with 25 ml lysis buffer [pH 8, 50 mM NaH_2PO_4 (Merck), 300 mM NaCl (Acros Organics), 10 mM imidazole (Acros Organics)] with 30% (v/v) glycerol and 2 L of wash buffer [50 mM NaH_2PO_4 (Merck), 300 mM NaCl (Acros Organics), 20 mM imidazole (Acros Organics), pH 8, 30% (v/v) glycerol (Acros Organics)] at 4 °C. Two elution fractions were collected by addition of 1 ml elution buffer [50 mM NaH_2PO_4 (Merck), 300 mM NaCl (Acros Organics), 250 mM imidazole (Acros Organics), pH 8, 30% (v/v) glycerol (Acros Organics)] and stored at 4 °C. Purity of the purified proteins was checked by SDS-PAGE and the yield was determined using a Nanodrop®.

3.2.4.3. Fast Protein Liquid Chromatography (FPLC)

Large scale purification of the recombinant proteins was conducted on a 1 ml HisTrap HP column (GE Healthcare) in combination with an Akta FPLC-system (GE Healthcare) and UNICORNTM 5.01 software. After column equilibration with 10 column volumes washing buffer [20mM NaH_2PO_4 pH 7.5, 0.5 M NaCl, protein specific imidazole (Acros Organics) concentration, 10% glycerol], the protein lysate was loaded onto the column. Next, the column was washed with 15 column volumes of washing buffer and subsequently eluted with elution buffer (20 mM NaH_2PO_4 pH 7.5, 0.5 M NaCl, 0.5 M imidazole, 10% glycerol).

3.2.4.4. Size-exclusion chromatography

Further recombinant protein purification was conducted with the size-exclusion chromatography (SEC) in combination with the Äkta System. Depends on recombinant protein size and its concentration after affinity chromatography purification, three different columns were used: HiLoad 16/600 Superdex 75 Prep grade gel filtration column, the HiLoad 16/600 Superdex 200 Prep grade gel filtration column and Superdex 200HP GL10/30 gel filtration column (GE Healthcare). The mobile phase buffer [50 mM Tris, 150 mM NaCl (Acros Organics), pH 7.6] was run through the column and

subsequently recombinant protein were eluted based on their molecular weight. Columns were calibrated by the LMW Calibration Kit (GE Healthcare).

3.2.4.5. Buffer exchange

Elution buffer was exchanged for PBS (for putative EPS depolymerases) or 5 mM HEPES (Acros Organics)/NaOH (UCB), pH 7.4 (for putative PG hydrolsases) supplemented with 30% (v/v) glycerol (Acros Organics) to improve stability and provide an appropriate medium for activity. Therefore, Amicon Ultra-0.5 mL Centrifugal Filters (10K and 30K, Merck) were used according to manufacturer protocol. Protein concentrations were determined by Nanodrop® ND-1000 spectrophotometer).

3.2.5. Functional assays

3.2.5.1. Halo formation plate assay

To get an indication of the potential polysaccharide degrading activity of putative depolymerases, a spot test was performed. *P. aeruginosa* PAO1 strain Krylov cells were grown overnight in LB broth and refreshed till OD_{600} ± 0.4. Next, 200 µl of this culture was added into 5 ml of 0.7% (w/v) soft agar and the mixture was poured on LB agar. Further, 10 µl protein and 10 µl negative control (either elution buffer or PBS) were spotted and incubated at 37°C. Plates were observed for halo formation after 24 h, 48 h and 72 h.

3.2.5.2. EPS staining and microscopy

P. aeruginosa PAO1 Krylov cells were removed from halo and bacterial zones created on LB agar. Cells were resuspended in a 10 µl droplet of 1% (w/v) aqueous Congo red solution (Sigma-Aldrich) on a microscopic glass slide. The suspension was spread over the entire slide and air-dried. A 10 µl drop of Maneval's solution [3.33% (v/v) phenol (VWR International), 4.44% (v/v) acetic acid (VWR International), 2.67% (w/v) Iron (III) chloride (UCB), 0.02% (w/v) fuchsin (Sigma-Aldrich)] was added and spread over the slide to acidify the Congo red background. Cells were observed under a light microscope at 100x magnification with immersion oil (Leica DM LB microscope; Mc Bain Instruments).

3.2.5.3. Biofilm degradation assay using the Calgary device

P. aeruginosa PAO1 Krylov cells were grown overnight in LB broth. Next, 1:200 dilution of overnight culture was inoculated in each well of a 96-well round-bottom plate.

The plate was covered with a peg lid and incubated at 37°C for 24 h, 48 h or 72 h without shaking. For 48 h and 72 h biofilms, the LB broth was refreshed every day. Next, the peg lids were rinsed with PBS buffer to remove planktonic cells and exposed to negative control (190 µl of LB broth), positive control (400 µg/ml gentamicin) or protein. Biofilms were further incubated at 37°C for 30 min without shaking. Next, the peg lids were washed with PBS buffer to remove planktonic cells and transferred to a 96-well flat-bottom plate containing 99% methanol (Chem-Lab) to fix the biofilm cells. After 15 min incubation at room temperature, the peg lids were air-dried and incubated for further 15 min with 0.01% (w/v) crystal violet (Merck). Next, biofilms were washed with tap water and air-dried. Finally, 33% (v/v) acetic acid (VWR International) was introduced for 5 min to dissolve the crystal violet and OD_{600} was measured using a Multiskan EX Microplate Reader (Thermo Scientific).

3.2.5.4. Activity assay of endolysins on permeabilized bacterial cells

To quantify the enzymatic activity of phage endolysins, the drop in OD_{655} of OM permeabilized *P. aeruginosa* PAO1, APEC or *Salmonella* Enteritidis substrate was monitored over time. Protoplasts of each strain were prepared by treating cells with chloroform-saturated Tris buffer according to the standard protocol of Lavigne *et al.* (2004). With this treatment, the protective effect of the OM is destroyed and the EAD of PG hydrolases can easily reach its target in the PG. Briefly, 20 ml overnight culture was inoculated in 500 ml of LB broth and grown to the mid-exponential phase (OD_{600} ± 0.6). Next, the culture was centrifuged (3220 × *g*, 30 min, 4 °C) and the pellet was resuspended in 240 ml of chloroform-saturated 0.05 M Tris buffer [chloroform (Merck), Trizma base (Acros Organics), pH 7.7]. After 45 min shaking at room temperature, the suspension was again centrifuged (3220 × *g*, 30 min, 4 °C) and the pellet was washed with 90 ml PBS. Finally, pellets were resuspended in 5 mM HEPES/NaOH (pH 7.4) to an OD_{600} of 1.5 and substrates were stored at -20°C. Prior to the assay, the substrate was thawed and sonicated to reduce the substrate viscosity and increase the homogeneity.

A volume of 30 µl of protein or its dilutions was mixed with 270 µl of substrate in a flat-bottom 96-well plate and the OD_{655} was measured over time using a Model 680 Microplate Reader (Bio-Rad). Positive controls (φKZ gp144 endolysin) and negative controls (5 mM HEPES/NaOH, pH 7.4) were included in the assay. The experiment was performed in triplicate and data were analyzed as described in Briers *et al*. (2007).

3.2.5.5. pH stability analysis of the endolysins

To assess the pH-dependency of PG hydrolases, the *P. aeruginosa* PAO1 substrate prepared as described in section 3.2.5.4 and resuspended in universal buffer adjusted to pH range 3-10 [150 mM KCl (Janssen Chimica), 10 mM KH_2PO_4 (Acros Organics), 10 mM sodium citrate dehydrate (Sigma-Aldrich) and 10 mM H_3BO_4 (Acros Organics)]. The experiment was performed in triplicate as previously described. The highest average slope, corresponding to a certain pH, was normalized and the relative activity of the other samples was calculated in percentage with standard deviation.

3.2.5.6. Zymography with crude peptidoglycan

The PG-degrading capacity of the purified proteins was assessed by zymography. *P. aeruginosa* PAO1 Krylov cells were inoculated in 500 ml LB broth and incubated overnight at 37 °C. Next, the culture was autoclaved and centrifuged (4570 × *g*, 30 min, 4 °C). Further, 300-600 mg of crude PG was copolymerized in a 15% or 12% (v/v) polyacrylamide separation gel. Additionally, a classical 15% or 12% (v/v) separation gel was prepared for comparison afterwards. A volume of 10 µl of purified protein samples were loaded on both gels, as well as 7 µl PageRuler™ Prestained Protein Ladder (Thermo Scientific) and two concentrations (10 µg/ml, 5 µg/ml) of HEWL (Sigma-Aldrich) as a positive control. SDS-PAGE was performed to separate the proteins according to their MW. Next, the gel was incubated for 48 h in renaturation buffer [150 mM NaH_2PO_4 pH 7 (Merck), 10 mM $MgCl_2$ (Sigma-Aldrich), 0.1% (w/v) TritonX-100 (Sigma-Aldrich)] and 24 h in distilled water to allow renaturation of the proteins and regeneration of the enzymatic activity. Finally gels were stained with zymography staining solution [0.10% (w/v) methylene blue (ICN Biomedicals), 0.001% (w/v) KOH (UCB)].

3.2.5.7. Antibacterial activity of endolysins

P. aeruginosa PAO1 Pirnay, APEC and *S.* Enteritidis cells were grown in LB broth to the mid-exponential phase (OD600 ± 0.6). Then, cultures were diluted 200-fold in 5 mM HEPES/NaOH (pH 7.4) to a final density of 10^6 cfu/ml. A volume of 100 µl of diluted cells were mixed with 50 µl of protein (2.5 µM final concentration) and 50 µl 5 mM HEPES/NaOH (pH 7.4) or 50 µl 2 mM Na_2-EDTA (Acros Organics)

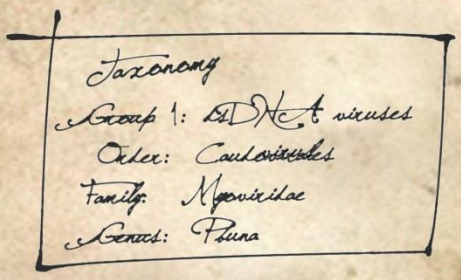
in 5 mM HEPES/NaOH (pH 7.4) (final concentration 0.5 mM). As controls, cells were incubated with only 2 mM Na_2-EDTA in 5 mM HEPES/NaOH (pH 7.4) or 5 mM HEPES/NaOH (pH 7.4). After 30 min of incubation at room temperature, mixtures were 10-, 100- and 1000-fold diluted in PBS and 100 µl of each dilution was plated on LB agar. The antibacterial activity of endolysins was quantified after 18 h incubation at 37 °C as the relative inactivation in logarithmic units (= $\log10(N_0/N_i)$ with N_0= number of untreated cells and N_i = number of treated cells counted after incubation). All experiments were performed in triplicate.

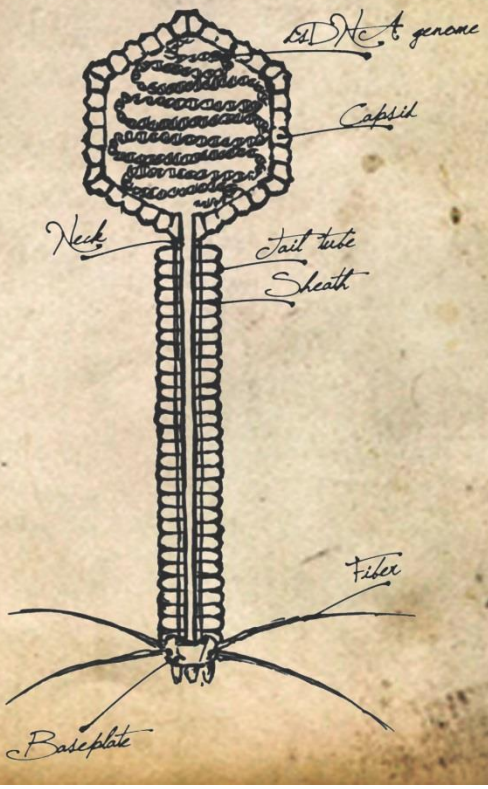
3.2.5.8. Time-lapse microscopy to visualize enzymatic activity

To visualize the mode of action of the proteins, a time-lapse microscopy was performed. *P. aeruginosa* PAO1 Pirnay cells were grown to mid-exponential phase ($OD_{600} \pm 0.6$) and centrifuged (4570 × g, 10 min, 4°C). Cell pellet was washed three times with PBS. Finally, cells were concentrated five times by resuspension in PBS buffer containing 0.5 mM Na_2-EDTA (Acros Organics). Equal amounts of protein and bacterial cells were mixed and fixed on a PBS-0.2% agar pad for observation under a temperature controlled Eclipse Ti-E inverted microscope (Nikon) equipped with a Ti-CT-E motorized condenser and a CoolSNAP HQ2 FireWire CCD camera, available at the Laboratory of Food Microbiology (KULeuven; Prof. A. Aertsen), with the help of Dr. Sander Govers. Images were recorded using NIS-Elements (Nikon) and further processed with ImageJ (http://rsbweb.nih.gov/ij/).

Chapter 4

Bacteriophages KTN6 and KTE28 represent the widespread and conserved Pbunavirus genus





Contribution

- Transmission electron microscopy phage visualisation and morphology analysis was performed by Prof. Hans-W. Ackermann (Department of Microbiology-Infectiology and Immunology, Medical School, Laval University, Quebec, Canada) and technician of University of Wrocław Sylwia Nowak.
- Time-lapse microscopy experiment was conducted in the Centre for Food and Microbial Technology (Prof. Abram Aertsen, KU Leuven, Belgium) with a help of Dr. William Cenens.
- LPS assay was performed by Dr. Daria Augustyniak (Department of Pathogen Biology and Immunology, University of Wrocław, Poland).
- Protein-sharing network was constructed and analyzed by Dr. Ho Bin Jang (Laboratory of Gene Technology, KU Leuven, Belgium) and Katarzyna Danis-Włodarczyk.
- Genome analysis was performed by Katarzyna Danis-Włodarczyk under supervision of Dr. Dieter Vandenheuvel.
- All remaining experiments and analysis were performed by Katarzyna Danis-Włodarczyk.

4.1. Introduction

The genus of *Pbunavirus* is a well characterized and widespread clade of closely-related phages, belonging to the *Myoviridae* family. Their genomes are highly conserved (more than 85% of nucleotide identity) and organized in at least seven transcriptional blocks. Their size varies between 64,427 and 66,530 bp with 88 to 95 encoded proteins (Ceyssens *et al.*, 2009). Phages from this group do not encode a recognizable integrase, suggesting their obligatory virulent nature (Lavigne *et al.*, 2009). They have a solid, acid resistant isometric capsid, with a diameter of ~74 nm and a contractile tail of ~140 nm (Ceyssens *et al.*, 2009). The LPS serves as the host receptor (Jarrell *et al.*, 1977). Over the years, no less than 44 phages have been reported to be *Pbunavirus*, mainly based on cross-DNA hybridization and morphological studies (Garbe *et al.*, 2010). Phages from this genus are described as a suitable candidates for therapy, especially when applied in phage cocktails (Merabishvili *et al.*, 2009). In this chapter, we describe two new broad host range *Pbunavirus* phages KT28 and KTN6, isolated from environmental water samples. Their genomes have been sequenced and a detailed comparison with other members of the *Pbunavirus* clade was made. The morphology, burst size, and host range were established as well as viral particle stability and LPS interactions.

4.2. Microbiological characteristics

4.2.1. Isolation and morphology

The KTN6 and KT28 phages were isolated from sewage samples collected in natural wastewater treatment plant (irrigated fields) located in Wrocław, Poland. Both lytic viruses propagate well on *P. aeruginosa* PAO1 reference strain. After purification phage titres were 10^9 - 10^{10} pfu/ml and caused 1.2–1.5 mm wide clear plaques on 0.6% soft agar with small halo zone, particularly evident in case a 167/DON clinical strain infection (Fig. 4.1). The bacteriophages were examined by transmission electron microscopy (TEM) and classified on the basis of their morphological features to *Pbunavirus*, family *Myoviridae*, order *Caudovirales* (Fig. 4.1). The isolates were formally named as vB_PaeM_KT28 (KT28) and vB_PaeM_KTN6 (KTN6). Their isometric head size was estimated to be 74 nm and 72 nm between opposite apices, and contractile tails 136 nm and 123 nm in the extended state, for KT28 and KTN6, respectively.

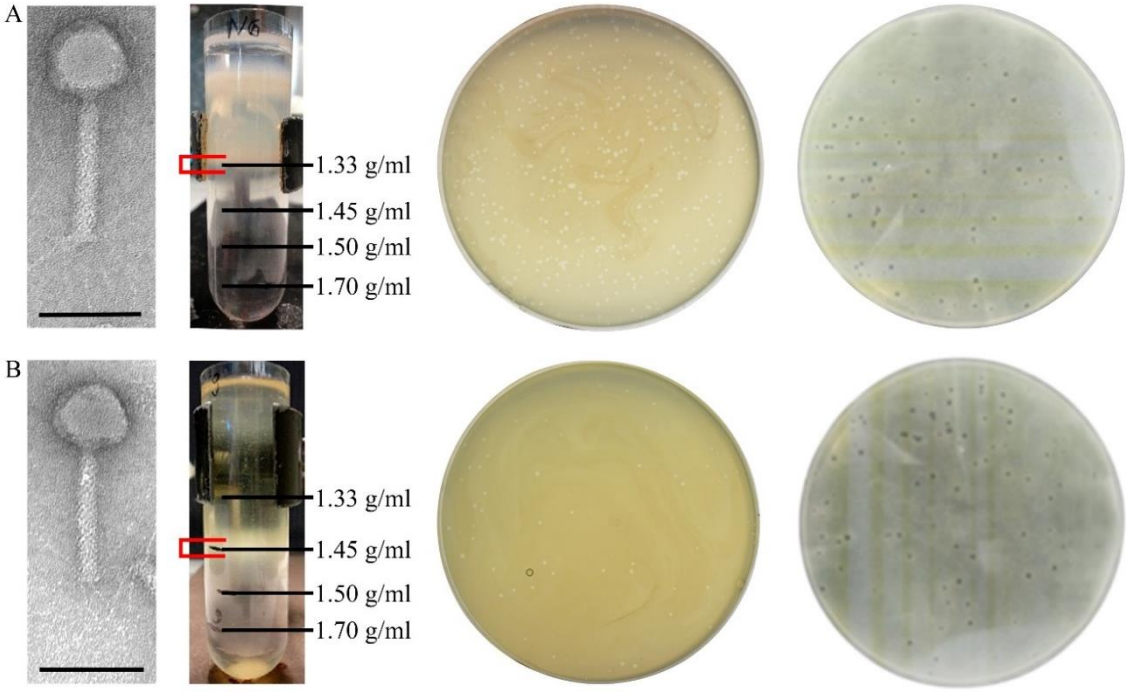


Figure 4.1: Particle analysis of the *Pseudomonas* phages KT28 (A) and KTN6 (B) belonging to *Pbunavirus* genus. From left to right, an electron microscopic image, a CsCl purification with phage band marked red, plaque morphology on *P. aeruginosa* PAO1 Krylov reference strain and 167/DON clinical isolate from the Academic Teaching Hospital in Wrocław, Poland. The scale bar represents 100 nm.

4.2.2. Phage infection

One-step growth experiments, constructed under standard laboratory conditions, indicated a latent period of 35 min and a burst size of about 64 ± 0.98 and 96 ± 0.2 phage particles per infected bacterial cell (pfu/cell), for KT28 and KTN6, respectively (Fig 4.2).

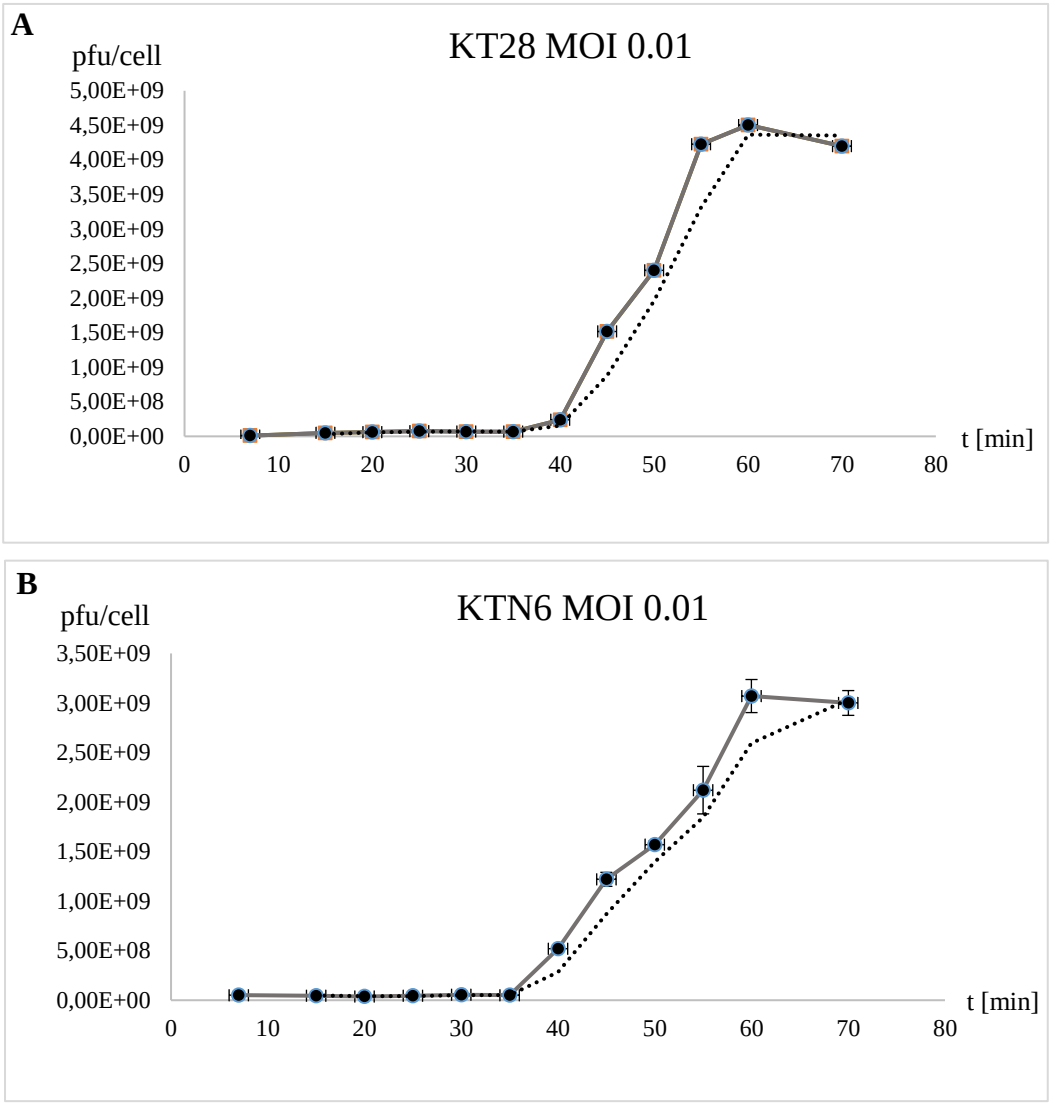


Figure 4.2: The one-step growth curves of KT28 and KTN6 phages. An equal volume of bacterial culture (at OD_{600} of 0.4, 10^8 cfu/ml) was mixed with phage suspension (10^6 pfu/ml) to obtain a multiplicity of infection (MOI) of 0.01. Phages were allowed to adsorb for 8 min at 37°C. Triplicate samples were taken during 70 min at 5 min intervals and titrated. The phage titre [pfu/cell] of the solution was assessed using the double-agar layer technique, according to methods described by Adams (1959). (A) KT28 and (B) KTN6 one-step growth curves with a latent period of 35 min and burst size of 64 ± 0.98 and 96 ± 0.2 phage particles per infected bacterial cell (pfu/cell), respectively. A trend line is marked with blue dots. All data are collected from three independent experiments, error bars indicate the standard deviation.

4.2.3. Time-lapse microscopy

The *P. aeruginosa* infection caused by phage KT28 or KTN6 in the real time was visualized with a time-lapse microscopy (Eclipse Ti Time-Lapse Microscope). *P. aeruginosa* PAO1 Krylov cells were grown to mid-exponential phase (OD_{600} of 0.6) and infected on a 0.2% PBS - agar pad by CsCl purified KT28 or KTN6 phages (6.5×10^{10} and 7.26×10^{10} pfu/ml, respectively) with MOI 10. In the course of infection cells only marginally increases their size and between 10 and 30 min post infection host cells started to burst (Fig. 4.3). The latent period of both phages achieved in this experiment cannot be compared to one step growth experiments, due to different conditions.

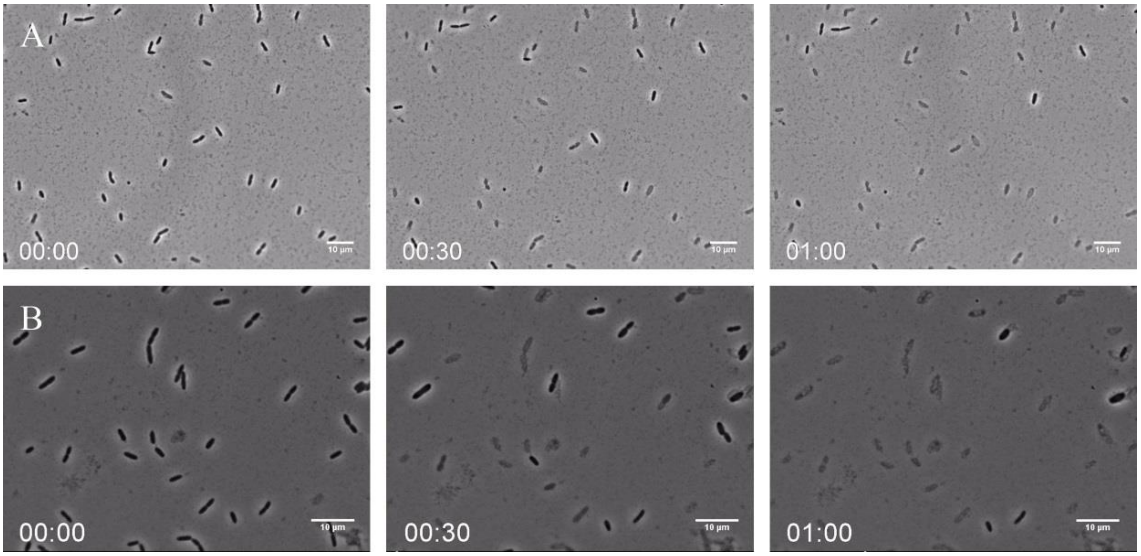


Figure 4.3: The time-lapse recordings of *P. aeruginosa* PAO1 Krylov cells infected with KT28 (A) or KTN6 (B) phages. A snapshots were taken from the start of phage infection, MOI 10, using a Nikon Eclipse Ti Time-Lapse Microscope. Scale bar represents 10 µm. In the left bottom corner is presented time in hours counted from the beginning infection.

4.2.4. Stability tests

The phage stability in the presence of physical and chemical agents was very similar for KT28 (1.01×10^9 pfu/ml) and KTN6 (1.26×10^9 pfu/ml).

No reduction of pfu/ml was noticed over a period of 60 min at a temperature range of 40 - 70°C and 5 min at 80°C (Fig. 4.4). The KT28 phage titre drops at 80°C by 2 logs after 15 min, 4 logs after 30 min, 5 logs after 45 min and 60 min of incubation (Fig. 4.4 A). The KTN6 phage titre drops at 80°C incubation by 2 logs after 15 min and 30 min, from 45 min phage titre drops to zero (Fig. 4.4 B). Both phages did not survived temperature of 90°C and 100°C.

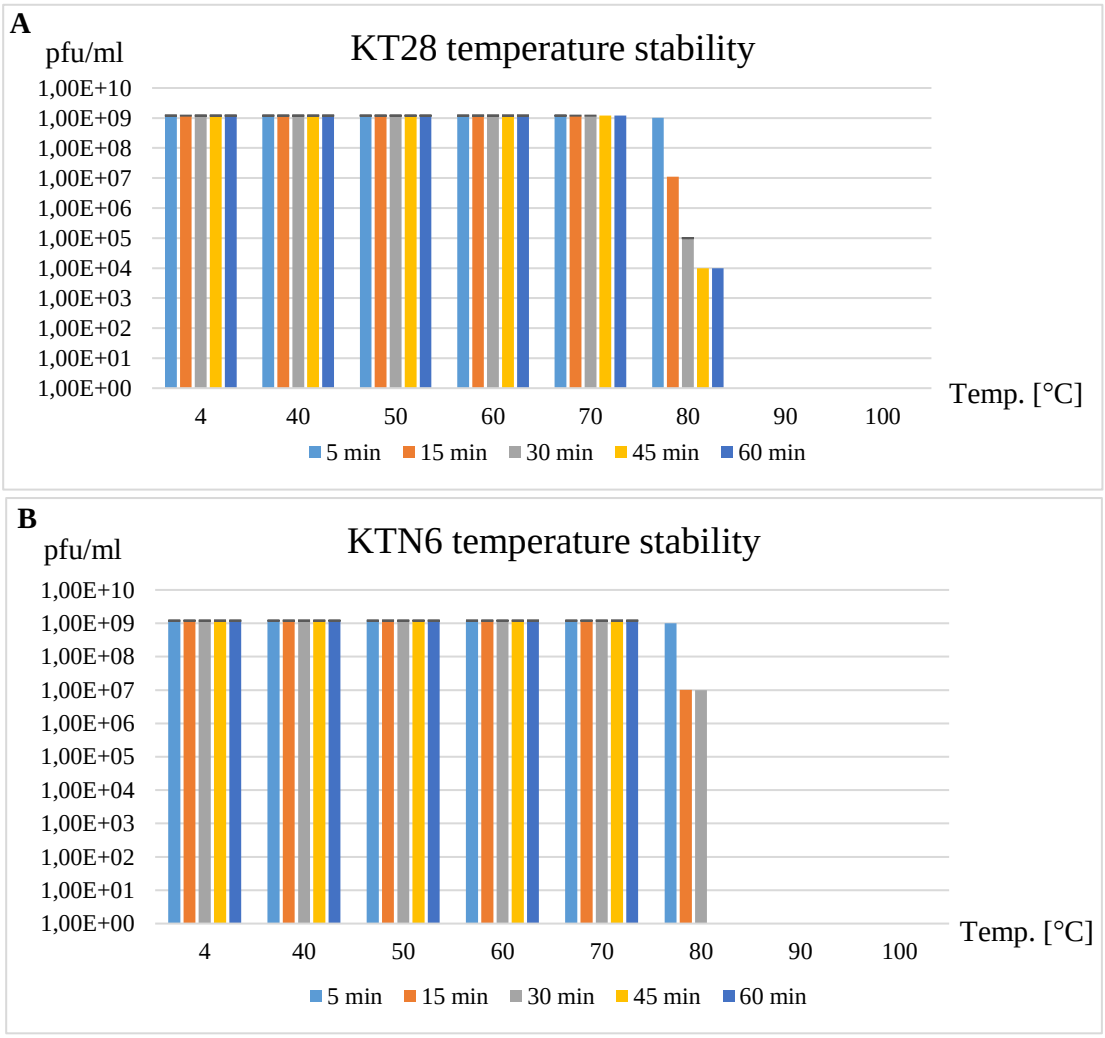


Figure 4.4: The temperature stability of phage KT28 (A) and KTN6 (B). Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

Phage particles were stable within a pH range of 3 - 12 after 1 hour incubation at room temperature. At pH 2 phage titre drops by 4 and 3 logs for KT28 and KTN6, respectively. At pH 1 both phages were not active (Fig. 4.5).

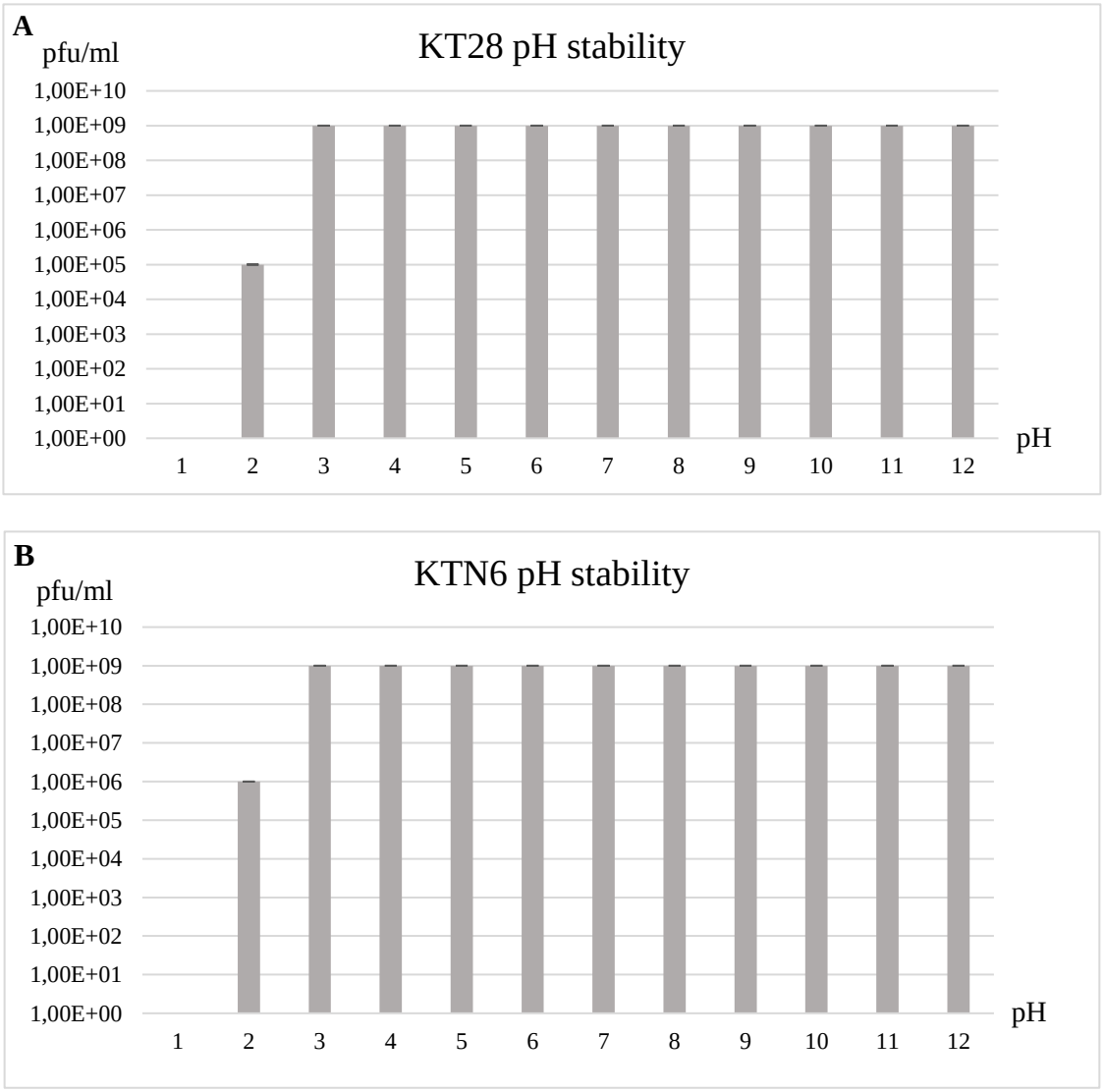


Figure 4.5: The pH stability of phages KT28 (A) and KTN6 (B). Phage titers (pfu/ml) were validated at the different pH values after 1h exposure at room temperature. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

Phage infectivity was not affected after 1 h and 24 h incubation with chloroform (an equal volume of phage solution and chloroform) at room temperature and at 4°C.

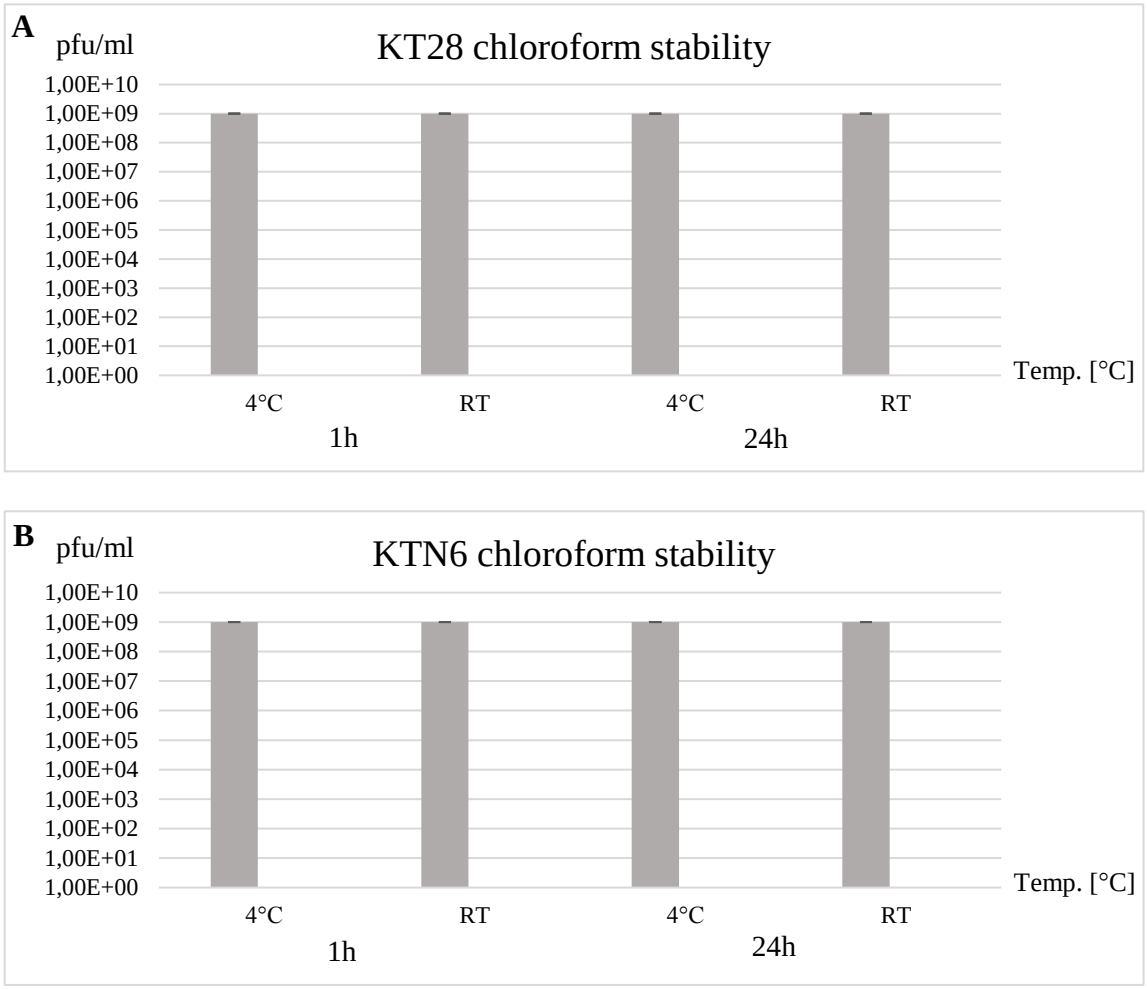


Figure 4.6: The chloroform stability of phage KT28 (A) and KTN6 (B). Phage titers (pfu/ml) were validated after chloroform treatment for 1 h and 24 h incubation at room temperature (RT) and at 4°C. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

4.2.5. Determination of phage receptor

Jarrell and Kropinski (1977) showed that phage E79 use the bacterial LPS layer as a receptor. In our study with knock-out PAO1 mutants, it was also confirmed that phages KT28 and KTN6 recognize the LPS elements as its specific receptor, and only absence of A- and B-band O-antigen inhibits phage propagation (Table 4.1).

Table 4.1. Host range analysis of KT28 and KTN6 phages with the use of *P. aeruginosa* PAO1 mutants.

Bacterial strain	Phenotype	Origin	KT28	KTN6
PAO1 (ATCC 15692)	Wild type	American Type Culture Collection	+	+
PAO1 Pirnay	Wild type with inactive type IV pili	Military Hospital Neder-Over-Heembeek, Brussels, Belgium	+	+
PAO1 Krylov	Wild type		+	+
PAO1 Δrmd (A-, B+)	Deficiency in D-rhamnose biosynthesis; lack of A-band LPS	Department of Molecular and Cellular Biology, University of Guelph, Canada	+	+
PAO1 ΔrmLC (A-, B-, core-)	Deficiency in L-rhamnose biosynthesis; truncate core region, lack of A-band and B-band LPS		-	-
PAO1 ΔwaaL (A-, B-)	Lack of WaaL ligating O-polymer to core-lipid A; LPS is devoid of A-band and B-band, semirough (SR-LPS, or core-plus-one O-antigen)		-	-
PAO1 ΔwbpL (A-, B-)	Lack of glucosyltransferase WbpL essential for initiation of both A-band and B-band synthesis		-	-
PAO1 ΔfliC ΔalgC ΔpilA	Lack of flagella; lack of AlgC required for A-band, core oligosaccharide, and alginate	Technical University Hamburg, Germany	-	-
PAO1 ΔfliC wt algC ΔpilA	Lack of flagella; lack of type IV pili		+	+
PAO1 ΔfliC wt algC wt pilA	Lack of flagella		+	+
PAO1 wt fliC wt algC wt pilA	Wild type		+	+

+ activity, bacterial lysis. - no activity.

These results were verified by the experiments performed with pure LPS isolated from PAO1 strain. The phage inactivation assay revealed direct correlation between LPS concentration and viral particles infectivity inhibition (Fig. 4.7), and that 36.0 μg/ml and 43.3 μg/ml of LPS was needed to inhibit the activity of 50% of 3 x 10^3 pfu, for KT28 and KTN6, respectively. The LPS-binding experiments based on biotinylated phages showed similar OD values for both phages (0.74 and 0.75).

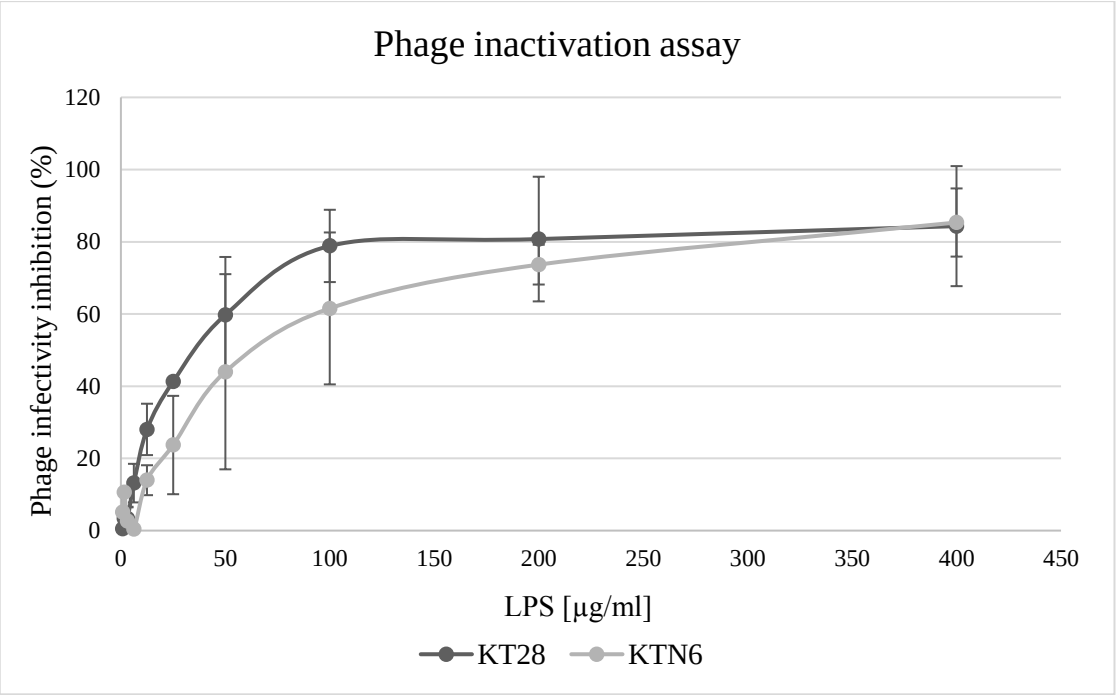


Figure 4.7: The KT28 and KTN6 phage inactivation assay by pure LPS. The curve represents direct correlation between LPS concentration (µg/ml) and viral particles infectivity inhibition (%). The light grey curve represents KTN6 phage and dark grey KT28 phage. All data are collected from three independent experiments, error bars indicate the standard deviation.

4.2.6. Host range analysis

Phage typing showed, that phage KTN6 exhibited a broader spectrum of activity against tested *P. aeruginosa* strains, lysing 68%, compared to 60% for phage KT28. This finding was also confirmed during phage typing of an international *P. aeruginosa* panel collection described by de Soyza *et al.* (2013). The KTN6 and KT28 propagated on 42% and 28% of the panel strains, respectively. Moreover, both phages were more active compared to three *Pbunavirus* phages: LMA2 25%, LBL3 47% and LSL4 17% (Table S 4.1).

Among 57 clinical strains of the Pirnay's *P. aeruginosa* panel seven were resistant to all phages used (Aa249, Lo049, Lw1047, Br667, Bu004, PAO29, Br229), 10 were sensitive to all five *Pbunavirus* phages chosen for treatment (Bu007, US448, Br906, PhDW6, Br642, LMG2107, Aa246, Pr335, Is580, Be136), eight to only KT28 and KTN6 (LMG14083, Lo053, So099, Is579, C, C1, LMG14084, B0548, PA6), four to KTN6 alone (Li004, C17, C18, US447), 3 to KT28 alone (C19, Br735, So095), 2 to LSL4 alone (US449, Br680) and 4 to LBL3 alone (Lo050, LMB5031, C2, Mi151).

4.2.7. Phage influence on biofilm

To initially evaluate biofilm eradication ability of KT28 and KTN6 phages, the 24 h and 48 h old *P. aeruginosa* PAO1 biofilms were grown on 8-well slide microscopy chambers and infected with KT28 or KTN6 phage suspension (1.5 x 10^{10} pfu/ml and 1.25 x 10^{10} pfu/ml, respectively), and further incubated for 24 h at 37°C. Two controls were set: 1) negative control of biofilm treated with LB broth, 2) positive control of a high dose gentamicin treatment (400 μg/ml). The MIC of gentamicin for tested strain was 0.5 μg/ml, however biofilm forming bacteria present different susceptibility to antibiotics in compare to their planktonic counterparts, thus a high dose of gentamicin was chosen. Biofilms were stained with LIVE/DEAD *Bac*Light Bacterial Viability Kit and visualized by confocal laser scanning microscopy (Fig. 6.7). We observe that the biofilm negative control presented a dense structure, with a thickness increasing over the time. However, the structure of the 24 h old biofilm was degraded after gentamicin and both phage treatments. In case of the 48 h old biofilm, gentamicin had a strong effect, degrading large areas of biofilm. In contrast, the phages treatment caused only small tunnels in biofilm structure in comparison to the control. This study show only preliminary assessment of KTN6 and KT28 antibiofilm activity, which was further evaluated with more qualitative and quantitative techniques by Drulis-Kawa group (Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Poland) and collaborators (Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University, Kielce, Poland). Their results are published in Danis-Wlodarczyk *et al.*, 2015.

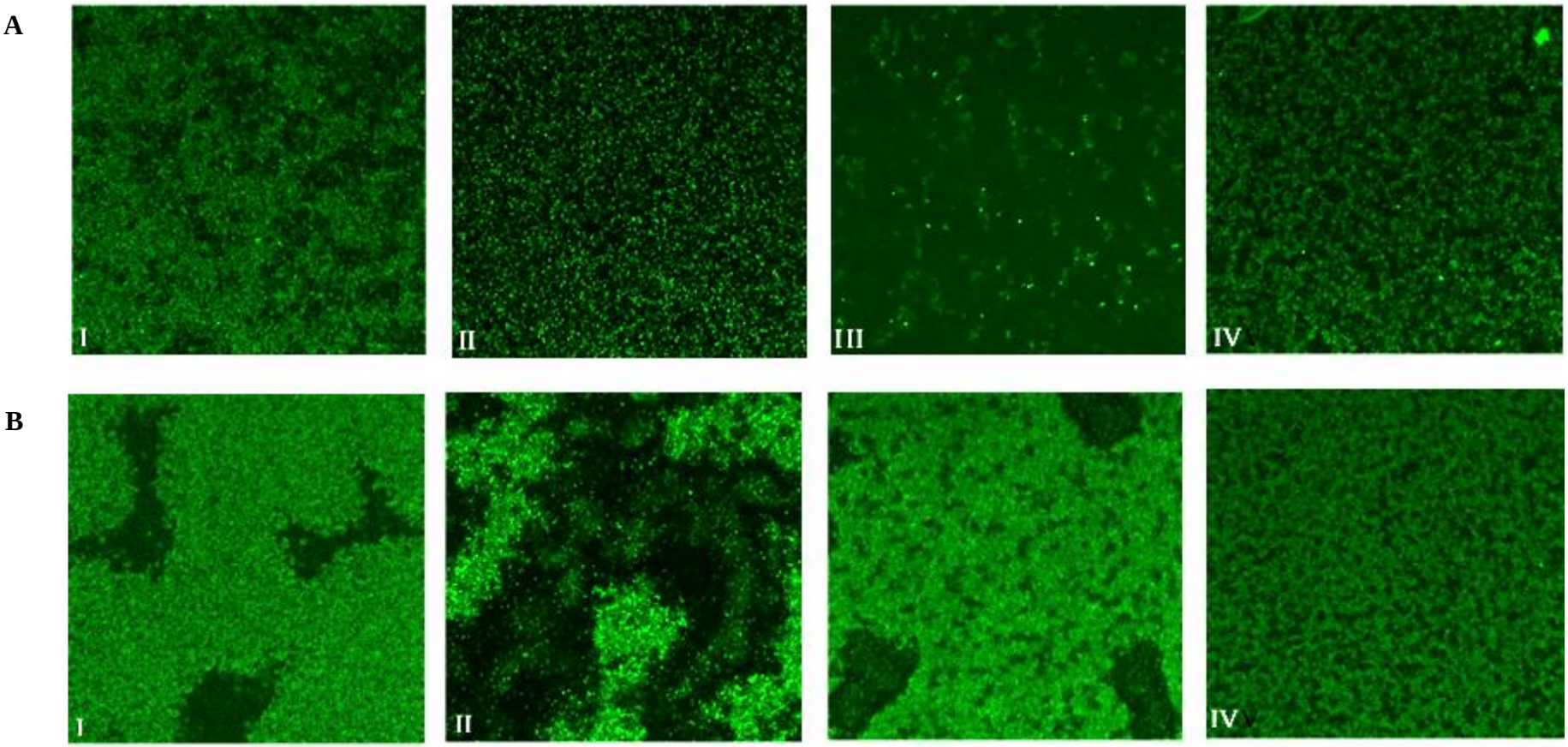


Figure 4.8: Antibiofilm activity of phage KT28 and KTN6 tested on *P. aeruginosa* PAO1 biofilms. (A) 24 h and (B) 48 h old biofilm. I) negative control: biofilm treated with LB broth; II) positive control: biofilm treated with gentamicin (400 µg/ml); III) KT28 treatment (1.5×10^{10} pfu/ml); IV) KTN6 treatment (1.25×10^{10} pfu/ml). Biofilms were grown at 37°C, 5% CO_2, stained with the LIVE/DEAD *Bac*Light stain (LIVE/DEAD *Bac*Light Bacterial Viability Kit) and visualized by confocal laser scanning microscopy (Zeiss LSM 510 Meta 40× objective, Jena, Germany). Live bacteria are stained green.

4.3. Genome analysis

4.3.1. Genome properties and annotation

Using high throughput sequencing by the Illumina MiSeq platform, the complete genome sequences of KTN6 and KT28 were determined (Fig. 4.9 and 4.10).

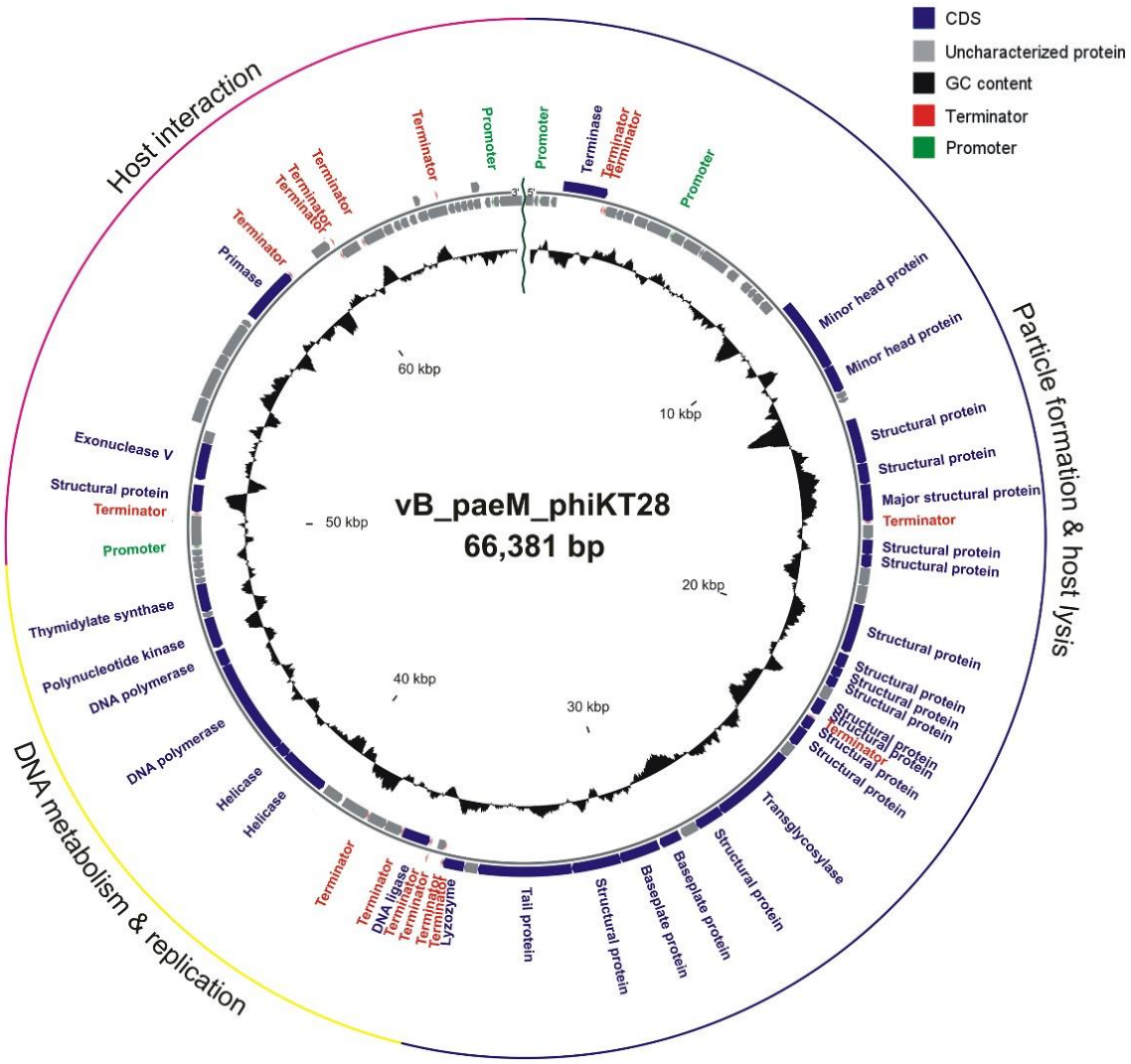


Figure 4.9: The circular representation of the KT28 genome. Blue arrows represent genes with predicted functions involved in DNA metabolism & replication, host infection and particle formation & host lysis. Grey arrows represent hypothetical proteins. Predicted (non-confirmed) promoters are labelled green while terminators red. The inner circle represents the GC content.

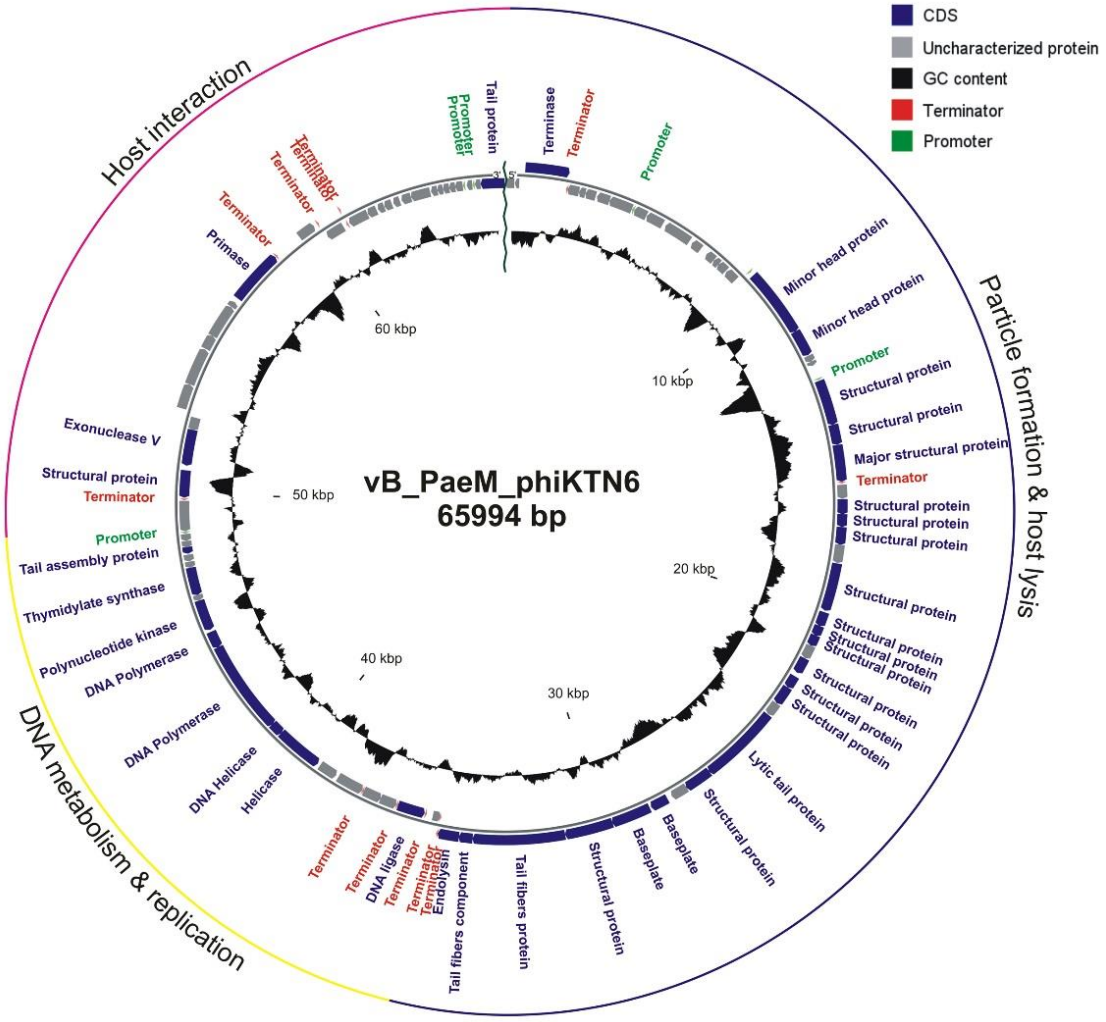


Figure 4.10: The circular representation of the KTN6 genome. Blue arrows represent genes with predicted functions involved in DNA metabolism & replication, host infection and particle formation & host lysis. Grey arrows represent hypothetical proteins. Predicted (non-confirmed) promoters are labelled green while terminators red. The inner circle represents the GC content.

Both viruses carry a non-permuted, linear double-stranded DNA genomes with GC content of 55.64% and 55.51%, respectively for KT28 and KTN6 (Table 4.2).

Table 4.2: General properties of phages KT28 and KTN6 and comparison to other *Pbunaviruses*.

Phage	Isolation details	Year	Genome size (%GC)	ORFs	Accession No.
KT28	irrigated fields, Wrocław, Poland	2011	66,381 bp (55.6)	96	KP340287
KTN6	irrigated fields, Wrocław, Poland	2011	65,994 bp (55.5)	91	KP340288
PB1	Sewage, Edinburgh, Scotland	1960	65,764 bp (55.5)	93	NC_011810
14-1	Sewage, Regensburg, Germany	2004	66,238 bp (55.6)	90	NC_011703
F8	Unknown	1972	66,015 bp (55.6)	93	NC_007810
LBL3	Pond, Blanes, Spain	2006	64,427 bp (55.5)	88	NC_011165
LMA2	River, Maastricht, Holland	2007	66,530 bp(55.6)	95	NC_011166
SN	Lake, Pokrov, Russia	2004	66,390 bp (55.6)	92	NC_011756

Phage KT28 has a genome size of 66,381 bp and KTN6 of 65,994 bp. In total, 94 and 91 open reading frames (ORFs) could be predicted, varying in size from 31 to 1036 and 46 to 1036 amino acid residues, respectively. Function was predicted for 32% (31 genes - KT28) and 40% (36 genes - KTN6) (Table S 4.2 and 4.3). The genomes have little intragenic space (6.96% - KT28 and 7.06% - KTN6) and 18 overlapping start/stop codons in case of both phages. Genes are organized in operons alternating on both strands. Furthermore, 4 and 3 bacterial promoters with AT-rich intergenic motifs were identified, for KT28 and KTN6, respectively (Fig. 4.11 A). For KTN6 additional 4 phage promoters were found, recognizable by the sigma factor, a bacterial transcription initiation factor, that enables specific binding of RNA polymerase to gene promoters.

B

KTN6

AAAATGCTTTACTTCCACTCTGGATGAAGGCAGAATAGCTCCATCGAAACGCGAAACCTTCAGACAAAAGG 4272-4342
AAAATGCTTTACTTCCACTCTGGATGAAGGCAGAATAGCTCCATCGAAACGCGAAACCTTCAGACAAAAGG 48503-48573
AAAGTGCTTTACTCGATGAGGTGACTTAGGTATAGTTACCTCATCGAAACGCGAAACACTCCAAAAAGAGG 64610-64680
GAAGTGCTTTACTCGATGAGATGACTGAGGTATAGTCATCTCATCGAAACCAGAAGCGCTCCAAAGAGGAA 64922-64992

Figure 4.11 Alignments of KT28 and KTN6 promoters. (A) Bacterial promoters prediction. The AT-rich intergenic motifs: -10 box and -35 box are marked with pink colour. The corresponding sequence logos from MEME/MAST are depicted below the alignments. On the left site are listed ORFs in front of which promoters are located, on the right side is sequence locus in the phage genome. (B) Phage promoters predicted for KTN6. On the right side is sequence locus in the phage genome.

Further, 15 and 12 rho-independent transcription terminators were predicted, respectively for KT28 and KTN6 (Fig. 4.12).

KT28

Gene	Strand	Motif	Sequence	dG (kcal/mol)	Location
ORF006	-	Both	GGCCGACATAATCACCGAGTCAGCGATGTTCGGTGAcTTGATCTTGCGT	-9.70	2519-2556
ORF006	-	Both	TAAGACCCGAAGCCCGGAGCGATCCGGGCTTTTCTATGTTTA	-15.00	2607-2637
ORF026	+	Both	CTATTCACCAAAGGCCGGGCTTCCGGCCTTTGTTCACTCTGA	-13.30	16424-16457
ORF037	+	Rnamotif	AATAGTTTCCGCCTCATCGTTCCCGATGGGGTTTTCCTTGTCG	-11.10	22161-22192
ORF037	+	Rnamotif	TTGAGGATAAGCCGGGTCAACACCCGGTTTTCTTCTGTA	-9.80	22713-22740
ORF050	-	Both	AGAAAGCAAAAAGGCCGCTGAATAAGCGGCCTTTTCTTTGGGCGA	-15.70	35704-35737
ORF049	+	Both	CCCAAAGAAAAGGCCGCTTATTCAGCGGCCTTTTTGCTTTCT	-13.80	35718-35748
ORF051	-	Erpin	GATCGAGATAAACCCGGACGCCTTATCCTACAGGCGATCCGGGTTTTCTCGTTTTAA	-9.71	36183-36228
ORF049	+	Both	AAAACGAGAAAACCCGGATCGCCTGTAGGATAAGGCGTCCGGGTTTATCTCGATCTA	-11.23	36195-36241
ORF052	-	Rnamotif	ACCGCCATTTAGCCCGGTTCGCCGGGCTTCTTTTCTGGA	-13.70	37174-37202
ORF054	-	Both	CACCGGCCGGAGAAAGGCGCTGTAATGGCGCCTTTCTTTTGAGGAGTC	-12.80	38381-38417
ORF069	-	Rnamotif	TAGCTCTTCAAGGGACTCATGAAAATGGGTCCCTTTTTATCCTCT	-13.20	50209-50242
ORF077	+	Rnamotif	TACTTGGACTAGGCCGCCTTCGGGCGGTCTTTTCTTTGGTG	-17.00	58525-58554
ORF079	-	Erpin	CCCAAAAGAAAGCCCCGGATTCTGAACTAGAGTCCGGGGCtTTGTCGTTTCTGA	-16.62	60141-60183
ORF078	+	Both	GAAACGACAAAGCCCCGGACTCTAGTTCAGAATCCGGGGCTTTCTTTTGGGCG	-16.12	60155-60196
ORF080	-	Rnamotif	CGCCCACCAAACACCGCAGCCTGTTCTGCGGTGTTTTCACATCCG	-13.80	60948-60981
ORF001	-	Rnamotif	GCTGGGACCCAGAGGCGGGCAAGTTTCTGCCTCgTTATAATTCCCA	-12.90	66134-66168

KTN6

Gene	Strand	Motif	Sequence	dG (kcal/mol)	Location
ORF004	-	Both	TAAGACCCGAAGCCCGGAGCGATCCGGGCTTTTCTATGTTTA	-15.00	2018-2048
ORF023	+	Both	CTATTCACCAAAGGCCGGGCTTCCGGCCTTTGTTCACTCTGA	-13.30	15836..15865
ORF047	-	Both	GAAAGCAAAAAGGCCGCTGAATAAGCGGCCTTTTCTTTTGGGC	-15.10	35114..35145
ORF046	+	Both	CCAAAAGAAAAGGCCGCTTATTCAGCGGCCTTTTTGCTTTCC	-13.80	35116..35157
ORF048	-	Erpin	GATCGAGATAAACCCGGACGCCTTATCCTACAGGCGATCCGGGTTTTCTCGTTTTAA	-9.71	35592..35636
ORF049	-	Rnamotif	ACCGCCATTTAGCCCGGTTCGCCGGGCTTCTTTTCTGGA	-13.70	36583..36610
ORF051	-	Both	CACCGGCCGGAGAAAGGCGCTGTAATGGCGCCTTTCTTTTGAGGAGTC	-12.80	37790..37826
ORF066	-	Both	TAGTTCTTCAAGGGACTCATGAGAATGGGTCCCTTTTTTATCCTT	-13.20	49613..49646
ORF074	+	Both	TACTTGGACTAGGCCGCCTTCGGGCGGTCTTTTCTTTGGCG	-17.00	58158..58187
ORF075	+	Both	AAACGACAAAGCCCGGACTCTAGTTCAGAATCCGGGGCTTTCTTTTGGGTG	-11.30	59793..59827
ORF077	-	Both	TCGCCAACAAACACCGCAGCCTGCTCTGCGGTGTTTTCACATCAGA	-13.80	60583..60617

Figure 4.12: The rho-independent transcription terminators in the genome of phage KT28 and KTN6. From the left site: list of ORFs after which terminators are located, strand and motif of terminator, terminator sequence with conserved stem-loop structures marked blue, dG (Δ*G*, the Gibbs free energy of stem-loop formation in kcal/mole) and location in phage genome.

Only two enzymes involved in nucleotide metabolism could by predicted: a thymidylate synthase (KT28 gp62 and KTN6 gp59) and a polynucleotide kinase (KT28 gp60 and KTN6 gp57). For DNA replication phages often use the DNA polymerase III enzyme, which consists alpha (KT28 gp58 and KTN6 gp 55) and epsilon (KT28 gp59 and KTN6 gp56) subunits of the catalytic core. The first subunit catalyses the polymerization reaction and have two conserved motifs ($_{342}$P**D**I**D**I**D**F and $_{492}$LL**K**I**D**A**LG**). Second is a DEDDh-type 3'-5' exonuclease important in the polymerase proofreading activity (Bernad *et al.*, 1989). These phages lack a phage-encoded RNA polymerase and they depend entirely on the host transcriptional machinery. Probably the replication elongation is independent of the host machinery, due to presence in both phage genomes phage-encoded DNA primase (KT28 gp77 and KTN6 gp74) and helicase (KT28 gp56, gp57 and KTN6 gp53, gp54).

The SDS-PAGE of CsCl purified KT28 and KTN6 virions presented a highly conserved banding pattern, which is consistent with the high sequence identity among *Pbunavirus* phages (Fig. 4.13). For both phages several structural proteins were predicted: two minor head proteins (KT28: gp20, gp21 and KTN6: gp17, gp18), 14 and 16 structural

proteins respectively, two baseplate related proteins (KT38: gp44, gp45 and KTN6: gp41, gp42), tail protein (KT28 gp47 and KTN6: gp44, gp38), tail fiber component (KTN6 gp45), tail assembly protein (KT28 gp65 and KTN6 gp62) and tail tape-measure protein (KT28 gp96 and KTN6 gp91).

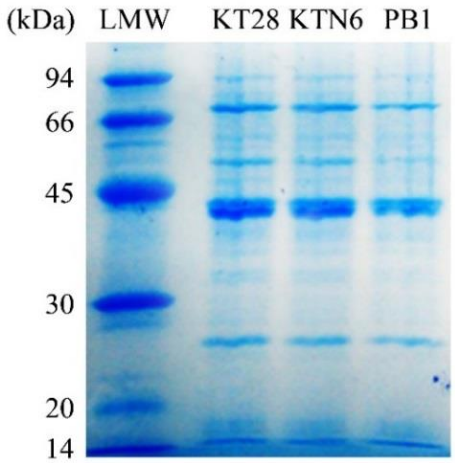


Figure 4.13: Phage particle proteins separated on a 12% SDS-PAGE gel. From left to right, a LMW molecular weight marker (kDa) and structural proteins of KT28, KTN6 and PB1 phages.

4.3.2. Comparative genome analysis and protein-sharing network

Phages show 96.4% nucleotide similarity to each other, but when compared to other *Pbunavirus* phages (14–1, F8, LBL3, LMA2, PB1, SN) they display nucleotide similarity between 47.7 – 96.2% for KT28 and 46.9 – 96.6% for KTN6 (Tab. 4.3), with most similarity to LMA2, 96.2% and 96.6%, respectively.

Table 4.3: The comparison of KT28 and KTN6 to sixn other *Pbunavirus* phages (14-1, F8, LBL3, LMA2, PB1 and SN) based on nucleotide similarity. Percent Identity Matrix created by Clustal2.1 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

Phage	14-1	F8	LBL3	LMA2	PB1	SN	KT28	KTN6
14-1	100	89	93	94	89	95	94	93
F8	89	100	91	89	94	89	89	90
LBL3	92	91	100	92	91	93	92	92
LMA2	93	89	92	100	88	95	96	97
PB1	89	94	91	88	100	88	88	89
SN	95	89	93	95	88	100	96	95
KT28	94	89	92	96	88	96	100	96
KTN6	93	90	92	97	89	95	96	100

When comparing *in silico* the protein identity of these new phages to PB1 phage, 71% of the proteins showed more than 90% similarity, especially in the regions responsible for particle formation & host lysis, and DNA metabolism & replication (Fig. 4.13). It confirms that the main core of *Pbunavirus* phage genomes is highly conserved. The largest protein variation was observed at the beginning and the end of the genomes.

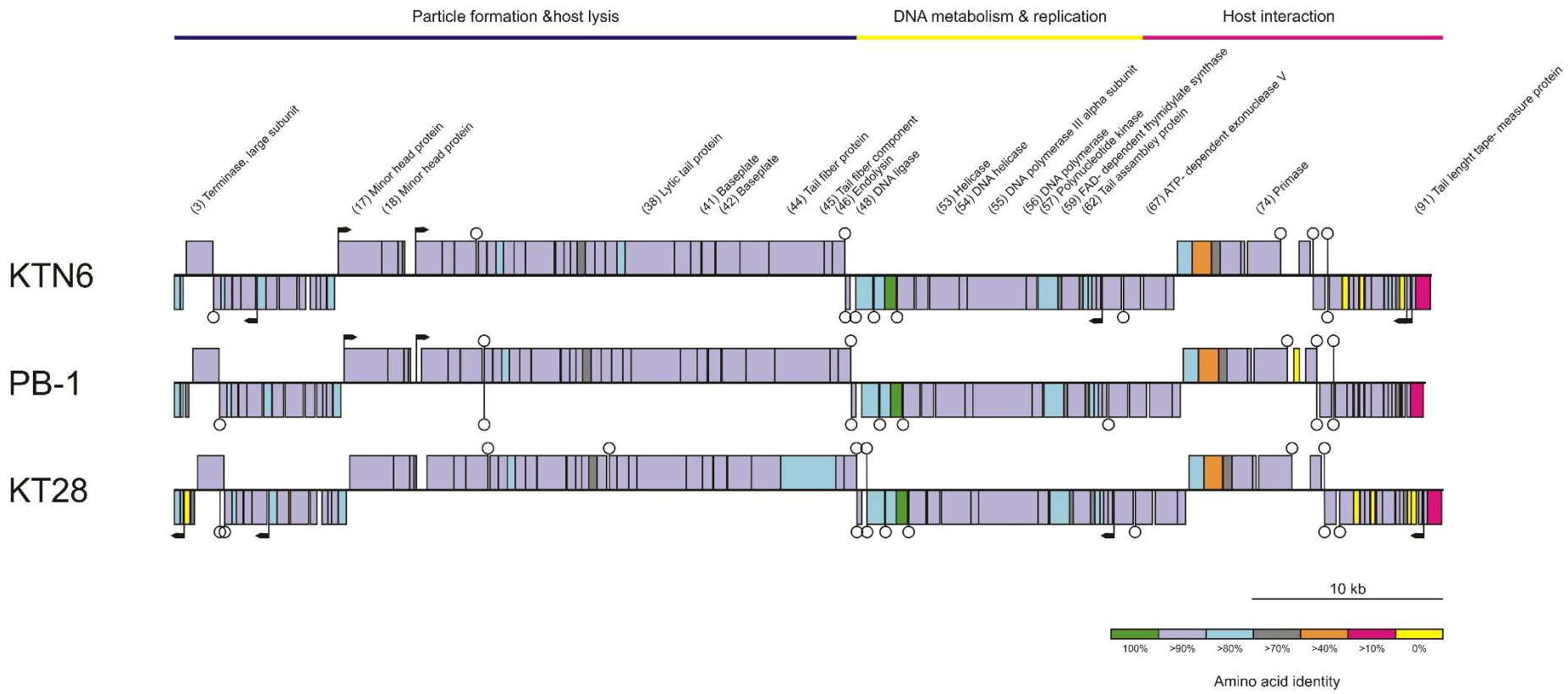
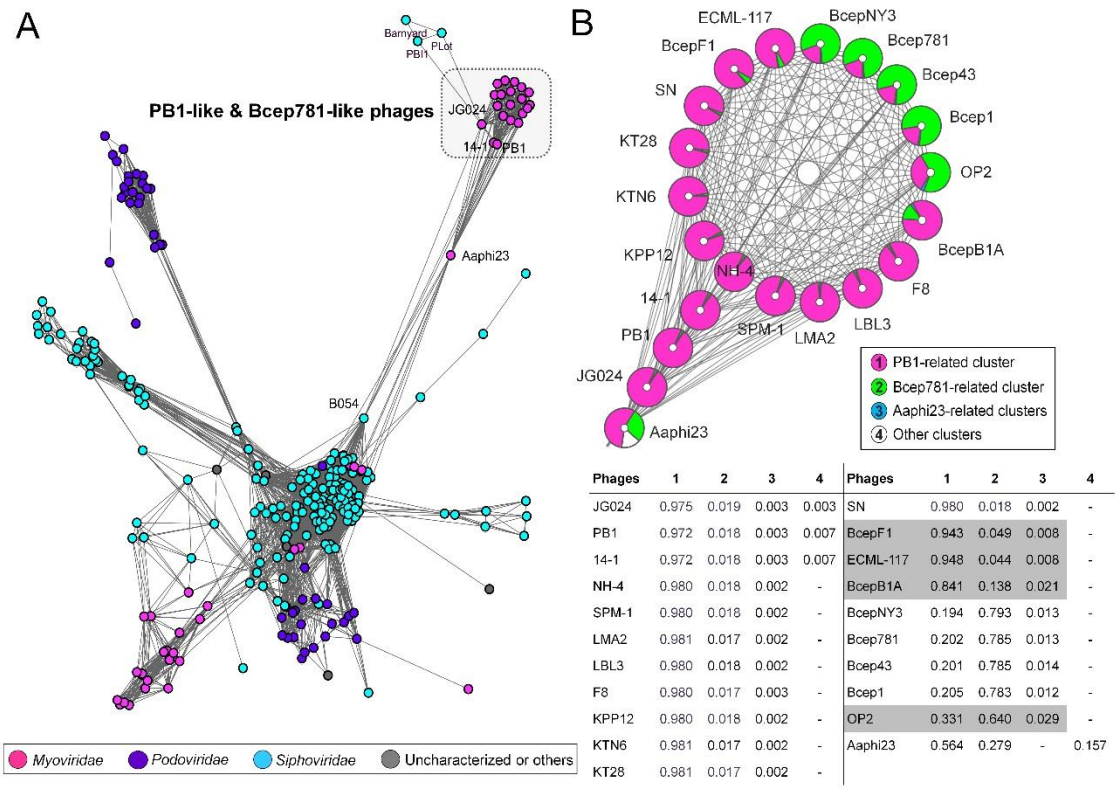


Figure 4.13: *In silico* analysis of KTN6 and KT28 phages and their comparison to PB1 phage. The predicted open reading frames of presented phages (rectangles) and their amino acid identity to the corresponding ORF in PB1 phage, indicated with different colors in the legend. Predicted terminators and promoters are shown as stem-loop structures and black arrows, respectively.

Figure 4.14 shows the resulting network consisting of 274 phages belonging to the *Myoviridae*, *Siphoviridae*, *Podoviridae*, or uncharacterized and another phages, and 4,928 relationships between them. In the network, KTN6, KT28, and other 13 phage members belonging to *Pbunavirus* group, together with *Bcep78virus* phages, were primarily placed into a small component at the periphery. Aaphi23 bridged such component comprising two phage groups with the largest component, whereas B054 linked to JG024, PB1, and 14–1; both of which are temperate phages (Dorscht *et al.*, 2009; Resch *et al.*, 2004). In addition, JG024 showed connections with the two mycobacteriophages PBI1 and PLot.



Phages	1	2	3	4
JG024	0.975	0.019	0.003	0.003
PB1	0.972	0.018	0.003	0.007
14-1	0.972	0.018	0.003	0.007
NH-4	0.980	0.018	0.002	-
SPM-1	0.980	0.018	0.002	-
LMA2	0.981	0.017	0.002	-
LBL3	0.980	0.018	0.002	-
F8	0.980	0.017	0.003	-
KPP12	0.980	0.018	0.002	-
KTN6	0.981	0.017	0.002	-
KT28	0.981	0.017	0.002	-

Phages	1	2	3	4
SN	0.980	0.018	0.002	-
BcepF1	0.943	0.049	0.008	-
ECML-117	0.948	0.044	0.008	-
BcepB1A	0.841	0.138	0.021	-
BcepNY3	0.194	0.793	0.013	-
Bcep781	0.202	0.785	0.013	-
Bcep43	0.201	0.785	0.014	-
Bcep1	0.205	0.783	0.012	-
OP2	0.331	0.640	0.029	-
Aaphi23	0.564	0.279	-	0.157

Figure 4.14: Protein-sharing network and reticulate relationship of *Pbunavirus*-related and *Bcep78virus*-related phage populations. (A) A network representation was produced using the edge-weighted spring embedded layout of Cytoscape version 3.1.1. Nodes indicate phages and edges between two nodes indicate their statistically weighted pairwise similarities with significance (Sig) scores of >1. In this way, 4,928 edges from 274 phages, including KTN6, KT28, and additional *Pbunavirus* phage members (JG024, PB1, 14–1, NH-4, SPM-1, LMA2, LBL3, KPP12, and SN) are shown and 159 phages that did not show any connections to those candidate phages were excluded for clarity. The node color corresponds to a given ICTV taxonomic classification of the *Myoviridae*, *Siphoviridae*, and *Podoviridae*, and uncharacterized and other phages. (B) Assigning of the phage clusters onto population network composed of interactions among *Pbunavirus* phages, *Bcep78virus* phages, Aaphi23, and others. Node pie charts show the proportion of homogeneity (membership) as estimated by estimating the proportion of the weight of connections to a given node to the nodes within a particular cluster. A color code for each cluster is represented in the legend box. Each membership value relevant to the represented phage is illustrated as the membership matrix in the lower panel.

To validate their positioning in the network, we used two classical population topologies such as the clustering coefficient (the extent to which the neighbors of a given node are interlinked) and the betweenness (a centrality estimating the proportion of shortest paths that pass through a node) (Assenov *et al.*, 2008). We found that with the highest clustering coefficient of 1 (1 means absolute cohesiveness) (Table 4.4), *Pbunavirus* phages including KTN6, KT28, F8, NH-4, SPM-1, LMA2, LBL3, KPP12, SN, ECML-117 and BcePB1A were tightly interconnected with *Bcep78virus* phages, implying their closest evolutionary relationship. However, compared to *Pbunavirus* members of a given component, JG024, PB1, and 14–1 showed lower clustering coefficients and higher betweenness distributions due to their more connected parts. Aaphi23 can be characterized by one of the highest betweenness node in the network.

Table 4.4: Clustering coefficient and betweenness centrality in the protein-sharing network of phage Aaphi23, and *Pbunavirus*/Bcep781-like phages.

Names	Clustering coefficient	Betweenness centrality
Aaphi23	0.56084656	0.10306764
JG024	0.76679842	0.03102882
141	0.91904762	0.0092124
PB1	0.91904762	0.0092124
KT28	1	0
KTN6	1	0
SN	1	0
BcepF1	1	0
OP2	1	0
Bcep781	1	0
KPP12	1	0
LBL3	1	0
BcepNY3	1	0
LMA2	1	0
Bcep1	1	0
Bcep43	1	0
F8	1	0
ECML	1	0
SPM1	1	0
NH-4	1	0
BcePB1A	1	0

Furthermore, we retrieved the reticulate clusters of *Pbunavirus* and *Bcep78virus* phages with the methodology used by Lima-Mendez *et al*. (2008), in which the weight of intracluster connection of a phage, called the membership, is used to differentiate cluster and membership > 0.79 can be a rough estimate of vertical evolution. In this reticulate classification, most phage members were highly assigned to their respective

ICTV genera with membership > 0.97 (Fig. 4.14). Notably, BcepF1, ECML-117, and BcePB1A showed lower memberships to PB1 cluster ranging from 0.841 to 0.948, which is consistent with their marginal relationships to *Pbunavirus* phages (Lavigne *et al.*, 2009; Carter *et al.*, 2012). Also, OP2 displayed a lower membership to Bcep781 cluster = 0.640 and thus more divergent among *Bcep78virus* members.

4.3.3. The lysis cassettes of KT28 and KTN6

Both analyzed phages have a globular endolysins, KT28 gp48 and KTN6 gp46, encoding lysozyme-like domains. A detailed description and experimental analysis is given in Chapter 8. Moreover, structural proteins KT28 gp41and KTN6 gp38 carry a transglycosylase SLT (soluble) domain (98.47% amino acid similarity, Clustal Omega) (Fig. 4.15), that degrade murein via cleavage of the beta-1,4-glycosidic bond between N-acetylmuramic acid and N-acetylglucosamine, with the concomitant formation of a 1,6-anhydrobond in the muramic acid residue. The muralytic activity of gp38 and gp41 was confirmed by use of a zymogram assay, showing specific degradation of peptidoglycan of autoclaved *P. aeruginosa* PAO1 cells upon phage protein renaturation (data shown in Chapter 8). These murein hydrolases assist the efficient DNA transfer to the host cell.

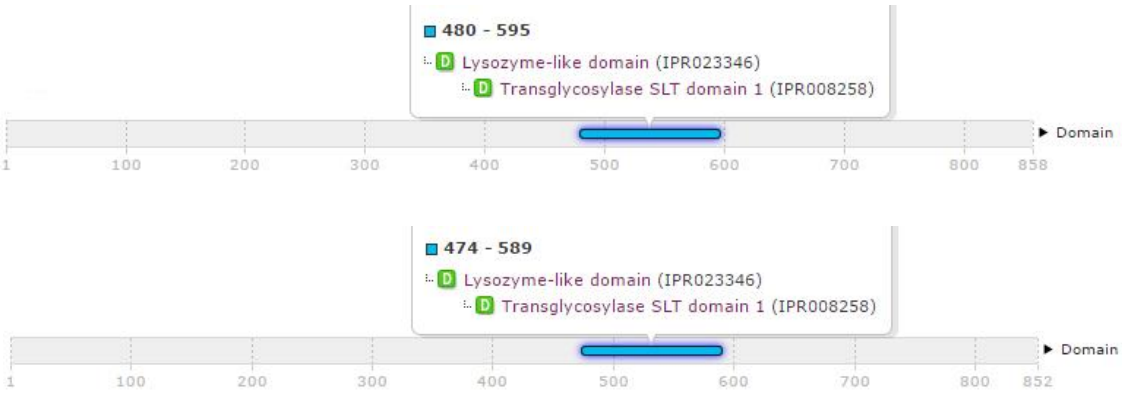


Figure 4.15: The KT28 gp41 (above) and KTN6 gp38 (below) transglycosylase domains predicted by InterPro (www.ebi.ac.uk/interpro/).

The general features of KT28 and KTN6 phages are summarize in table 4.5.

Table 4.5: Major features of KT28 and KTN6 phages.

Phage	KT28	KTN6
Water source	Irrigation field	
City/Country/ Date	Wrocław/Poland/ 2011	
Host	PAO1	
Genome size (bp)	66,381	65,994
GC (%)	55.64	55.51
ORFs	94	92
Unique ORFs	1	-
Virion dimensions head/tail [nm] [a]	74/136	72/123
Latent period [min]	35	35
Burst size [pfu/cell]	65	96
Heat stability for 1h [ºC]	40 - 70	40 - 70
pH stability for 1h at RT	3 - 12	3 - 12
Chloroform sensitivity at RT/4 ºC [h]	24	24
Phage infectivity inhibition by LPS (PI_{50}) [ug/ml] [a]	36.0	43.3
LPS-binding assay [OD 450nm][a]	0.74	0.75
Biofilm eradication [CFU reduction] [a]	70 - 80%	70 - 90%
Pyocyanin production inhibition in biofilm structure [b]	+	+
Pyoverdin production inhibition in biofilm structure [b]	+	+
Increase of diffusion though biofilm structure [b]	+	+
Biofilm eradication measured by goniometry analysis [b]	+	+

[a]Experiment performed by Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland (Danis-Wlodarczyk *et al.*, 2015).

[b]Experiments performed by Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University in Kielce, Kielce, Poland; Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland (Danis-Wlodarczyk *et al.*, 2015).

4.4. Discussion

Microbiological characteristics

The virions of newly isolated KT28 and KTN6 show characteristic features of *Pbunavirus* with respect to head/tail size and LPS host receptor recognition. The activity of phage virions was inhibited by 36.0 μg/ml and 43.3 μg/ml of pure LPS for KT28 and KTN6, respectively. The KT28, with a larger genome encapsulated in a bigger virion, exhibits the same latent period, but much lower burst size in comparison to KTN6 phage (64 ± 0.98 pfu/cell versus 96 ± 0.2 pfu/cell). Both phages possess similar susceptibility to chemical and physical agents such as high temperature, chloroform and pH. Phage infectivity tested on clinical strains revealed that both phages have broad host range. However, phage KTN6 was more potent than KT28, propagating on 68% versus

60% of the strains, respectively. The broad host range was also observed in case of other, previously described *Pbunaviruses*, such as LBL3, SN (Plenteneva *et al.*, 2009) or 14-1 (Merabishivili *et al.*, 2009).

Biofilm matrix degradation

The opportunistic pathogen *P. aeruginosa* has evolved a number of virulence features including: (i) biofilm formation, (ii) secretion of factors that impair neutrophil phagocytosis and activation, (iii) type III secretion system-dependent cytotoxicity (Dacheux *et al.*, 2000; Mahajan-Miklos *et al.*, 1999). Antibacterial therapies using these newly isolated phages may target all the above features to reduce the severity of infection and increase the efficacy of immune response in the bacteria eradication process. In this chapter we presented only preliminary evaluation of biofilm degradation ability of KT28 and KTN6 phages. Further experiments were conducted by the Drulis-Kawa group and collaborators on both peg-lid plate assay and PET membrane surface (Danis-Włodarczyk *et al.*, 2015) (data not shown). Their results showed that biomass evaluation by CV staining was not sensitive enough to detect any significant changes after phage treatment. Moreover, CV staining revealed unstable results with high standard deviations. As a consequence, other methods were used to evaluate phage potency to affect biofilm forming bacterial cells. For this purpose, a colony count technique was applied, the analysis of pyocyanin and pyoverdin secretion by spectrophotometry and fluorometry as well as laser interferometry and goniometry for the growth medium diffusion through biofilm matrix analysis. The standard CFU evaluation revealed a significant reduction (70% - 90%) in 24 - 72 h old biofilm cultures for both *Pbunavirus* phages treatments, suggesting that phage particles were able to penetrate the biofilm matrix. Both active phages showed also a strong inhibitory activity on *P. aeruginosa* biofilm, regardless of the bacteria origin (laboratory, clinical isolate from wound infection and clinical isolate from CF patient). Furthermore, phage KT28 was able to influence the amount of pyoverdin secreted by PAO1 strain and burn wound isolate 0038. The KTN6 phage affected only 72 h old culture in active form, but of all three tested strains. Moreover, the amounts of the most important dyes secreted by the *Pseudomonas* (pyocyanin and pyoverdin) decreased significantly, which may prove the potency of isolated phages to be applied as efficient antibacterials in *Pseudomonas* biofilm treatment. Further analysis, including diffusion and goniometry experiments, support this statement. The increase of diffusion rate through the biofilm matrix after phage application,

indicated the degradation of the biofilm structure associated with the loss of biofilm matrix elements on membrane surface and the reversion to hydrophobic properties of membrane, or partial degradation of matrix/biofilm structure. The increase of the diffusion was also obtain after the application of UV inactivated particles, suggesting that both tested phages KT28 and KTN6 are equipped with exopolysaccharide depolymerases responsible for phage particle spread within the biofilm matrix, however gene identification was difficult due to lack of homology in available databases.

Unfortunately, the biofilm treatment by these phages caused also the emergence of phage-resistant variants of *P. aeruginosa* from the biofilms. The persistent cells turned out to be cross resistant to both phages, probably by the loss or changes in O-antigen structure. This feature was relatively stable and persistent bacteria had a small colony variant morphology. It may put the applicability of *Pbunavirus* phages into question, although such enveloped modification may also decreases the fitness of the microbe and made them more susceptible to immune system defense.

Genome features

Using high throughput sequencing by the Illumina MiSeq platform, the complete genome sequences of KTN6 and KT28 phages were determined. The GC content (55.64% and 55.51%, respectively for KT28 and KTN6) is fairly constant throughout the genome and significantly lower than that of their host *P. aeruginosa* (66.6%). The genomes of both phages present high degree of DNA sequence similarity to the other *Pbunavirus* representatives. The biggest sequence variations occurred at the beginnings and at the ends of genomes. The ORFs are organized in operons situated on both strands and closely packed with little intergenic species (6.96% - KT28 and 7.06% - KTN6) and several overlapping start/stop codons.

Function was predicted for only 32% (31 genes-KT28) and 40% (36 genes-KTN6), including proteins involved in nucleotide metabolism & replication, host interaction and well as several structural proteins. Numerous small genes encoded proteins with unknown function. These phages lack phage-encoded RNA polymerases and are entirely dependent on host transcription, as was observed in other *Pbunaviruses*, such as LBL3, LMA2, 14-1 and SN (Ceyssens *et al.*, 2009). Phages from this clade have also a highly conserved proteome pattern, which is consistent with the high sequence identity among phages from this genus.

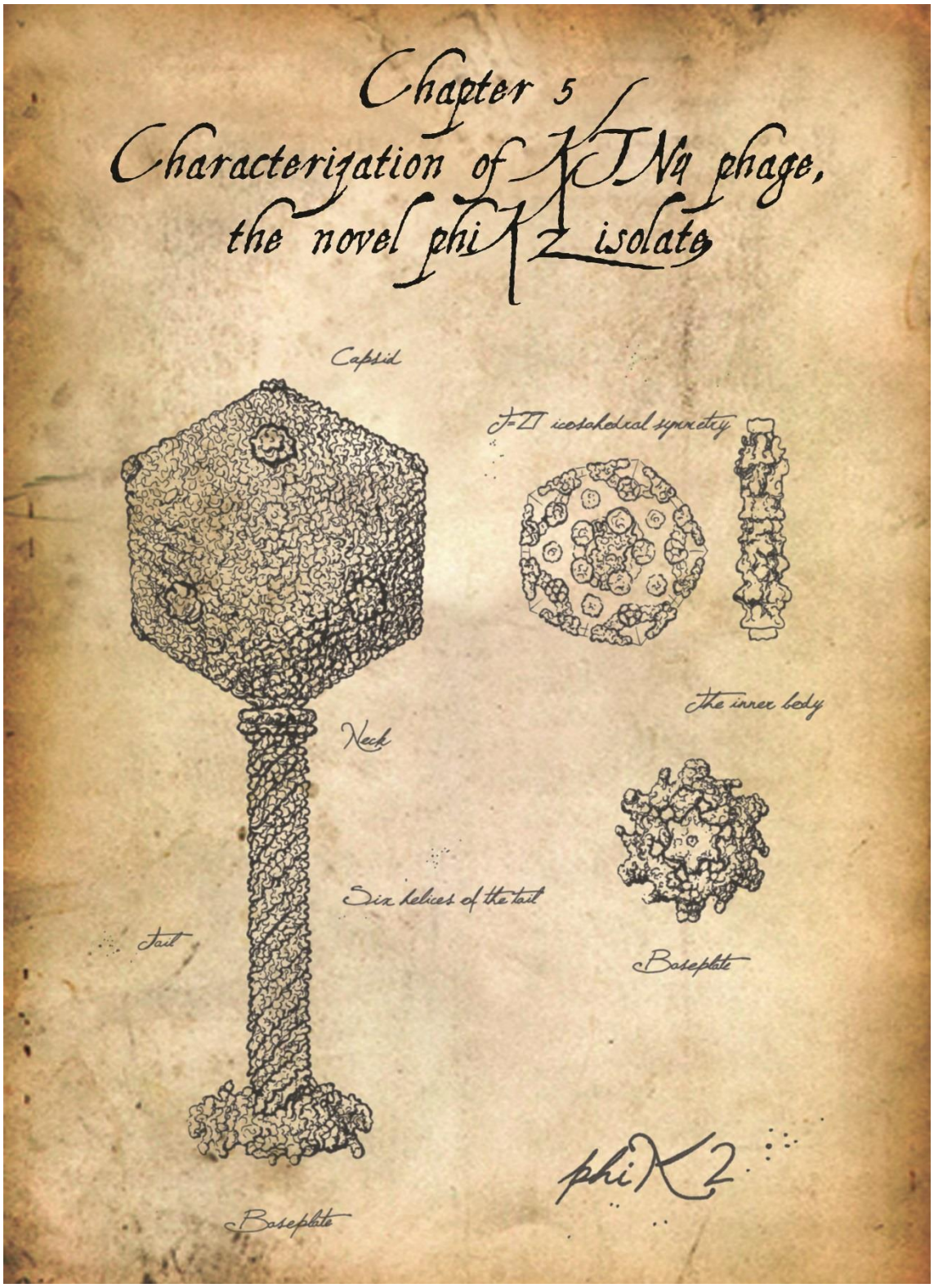
Furthermore, both phage genomes encode a globular endolysins, KT28 gp48 and KTN6 gp46 with lysozyme-like domains, and proteins with transglycosylase SLT domain (KT28 gp41 and KTN6 gp38), that have potential ability to degrade bacterial peptidoglycan.

Protein-sharing network

Genome comparison showed that KT28 and KTN6 are highly related to the widespread and conserved *Pbunavirus* phages. This group of viruses shows a high sequence similarity and limited horizontal gene transfer (Ceyssens *et al.*, 2008). To investigate genome evolution of KTN6 and KT28, we used a mathematical model of protein-sharing network, extending to probable close relatives. A graphical representation of the network can reveal several evolutionary connections of *Pbunavirus* -related genomes to other phage genomes. A major observation is that most of the *Pbunavirus* members were densely interconnected to those of the *Bcep781virus* group, as measured by their clustering coefficients. Such near-one clustering indicates the conservation of a core genome among those two groups (Lima-Mendez *et al.*, 2008). However, more importantly, they did not form a single group due to multiple connections of JG024, PB1, and 14–1 of the *Pbunavirus* group to the temperate phage B054 (Dorscht *et al.*, 2009) or two mycobacteriophages PBI1 and PLot. Based on this observation, two homologous genes were observed which only exist in all six corresponding phages and their positions in the limited conservation regions of the genomes of JG024, PB1, and 14–1 (Dagan, 2011); ORF 13 of JG024 and 14–1, and ORF 94 of JG024 and PB1, and ORF90 of 14–1. This may indicate that gene gain or loss in such variable regions could be important for their diversity, despite the limited lateral gene transfer in *Pbunavirus* group (Ceyssens *et al.*, 2009). Such interesting links are revealed through our evolutionary reconstruction of gene phylogenies (Dagan, 2011). In addition, another temperate phage Aaphi23, that shows some chimeric features (Lima-Mendez *et al.*, 2008; Resch *et al.*, 2004), acts as a bridge for both *Pbunavirus* and *Bcep781virus* groups. This reflects the tendency of temperate phage-mediated genetic exchange between strictly lytic phages, *Pbunavirus* and/or *Bcep781virus* phages, and the rest of the phage population (Lima-Mendez *et al.*, 2008; Lawrence *et al.*, 2002). Collectively, together with a reticulate classification, our phage population network revealed that *Pbunavirus* and *Bcep781virus* phages most likely have diverged through vertical evolution with some vertically

inherited modules and their ancestral relationship might be beyond the genus level at the subfamily level.

Chapter 5
Characterization of KTN4 phage, the novel phiKZ isolate



Contribution

- TEM phage visualisation and morphology analysis was performed by Hans-W. Ackermann (Department of Microbiology-Infectiology and Immunology, Medical School, Laval University, Quebec, Canada) and Sylwia Nowak, technician of University of Wrocław.
- Time-lapse microscopy experiment was conducted in the Centre for Food and Microbial Technology (Prof. Abram Aertsen, KU Leuven, Belgium) with a help of Dr. William Cenens.
- ASL and biofilm experiments were performed during research visit in Department of Molecular Medicine, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland (Prof. Brian J. Harvey) with the support of Dr. Gerard Higgins and Dr. Jean Tyrrell.
- The protein-sharing network was constructed and analyzed by Dr. Ho Bin Jang (Laboratory of Gene Technology, KU Leuven, Belgium) and Katarzyna Danis-Włodarczyk.
- ESI-MS/MS was conducted in Biomedical Research Institute and Transnationale Universiteit Limburg, School of Life Sciences, Hasselt University, Diepenbeek, Belgium (Prof. Jean-Paul. Noben).
- All remaining experiments and analyses were performed by Katarzyna Danis-Włodarczyk.

5.1. Introduction

To date, over twenty φKZ-related phages have been found in diverse geographic locations. They belong to a group of "giant" myoviruses and are infective for a variety of *Pseudomonas* species (Krylov *et al.*, 2007). Typically, they have a very large icosahedral head, ~122 nm in diameter, and a long (~190 nm) contractile tail surrounded by fibers. The capsid contains an inner body, which has been speculated to organize the packaged DNA (Fokine *et al.*, 2007; Krylov *et al.*, 1984). Furthermore, φKZ-related phages have an exceptionally large and complex circularly permuted, and terminally redundant linear dsDNA genomes (between 211 and 317 kb) (Mesyanzhinov *et al.*, 2002; Thomas *et al.*, 2008 and 2012). They display a pronounced difference in GC content (36.8 - 48%) and all have markedly lower GC content than the chromosomes of their GC-rich *Pseudomonas* hosts (60 to 66% GC) (Cornelissen *et al.*, 2012).

The φKZ-related phages have diverged strongly from the *Myoviridae* and present a higher divergence rate than *T4virus* genus or host genomes (Thomas *et al.*, 2008 and 2012). Therefore, a straightforward functional assignments of their encoded proteins has been seriously hampered even for the conserved core genes (e.g. those encoding tail fibers, baseplate module and DNA polymerase), as a result of low protein sequence similarity. Nevertheless, the φKZ-related phages and their gene products are important due to their potential to control pathogenic pseudomonads (Briers *et al.*, 2008;

Miroshnikov *et al*., 2006; ParadisBleau *et al*., 2007; Mesyanzhinov *et al*., 2002) and genetically manipulate their hosts (Monson *et al*., 2011). Representative of *Phikzvirus* clade, φKZ, have already been incorporated into traditional phage therapy cocktails and continue to be examined for novel therapeutic applications [e.g., (Golshahi *et al*., 2011; Matinkhoo *et al*., 2011)].

In this chapter we describe the morphology, genome organization, biology and antibacterial potential of a novel φKZ isolate KTN4, a potentially suitable candidate for *in vivo* phage therapy.

5.2. Microbiological characteristics

5.2.1. Isolation and morphology

Lytic phage KTN4 was isolated from sewage samples collected from irrigated fields located in Wrocław, Poland. After purification phage titres were 10^{10}–10^{11} pfu/ml and caused ~1.7 mm wide clear plaques with halo zone on 0.6% soft agar (Fig. 5.1). The KTN4 morphology was examined by TEM and classified to the *Phikzvirus*, order *Caudovirales*, family *Myoviridae* (Fig. 5.1). The isolate was formally named vB_PaeM_KTN4 (KTN4). The size of icosahedral head can be estimated at 130 nm between opposite apices, the contractile tail 168 nm long surrounded by fibers.

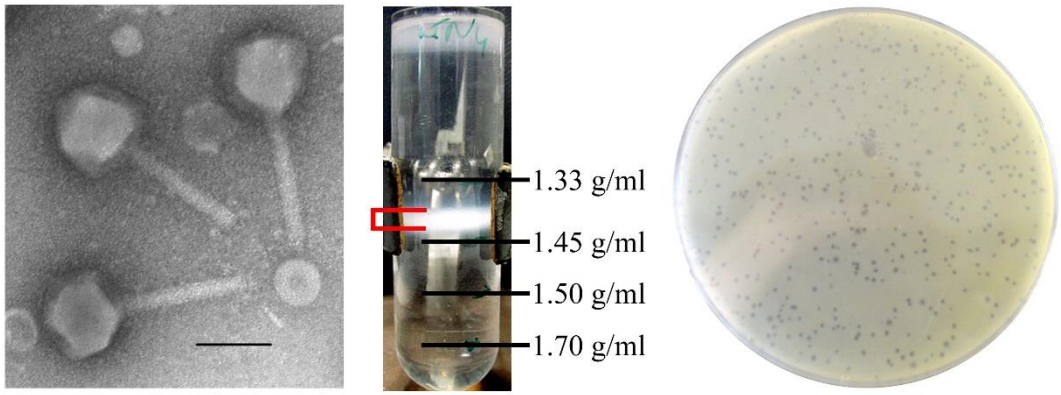


Figure 5.1: Particle analysis of the KTN4 phage. From left to right, an electron microscopic image with the scale bar representing 100 nm, a CsCl gradient purification (4 different CsCl densities are marked: 1.33 g/ml, 1.45 g/ml, 1.50 g/ml, 1.70 g/ml; phage band bracketed in red), plaque morphology on *P. aeruginosa* PAO1 Krylov reference strain.

5.2.2. Phage infection

One-step growth experiments indicated a latent period of 40 min and a burst size of about 8 ± 2 phage particles per infected bacterial cell (Fig. 5.2).

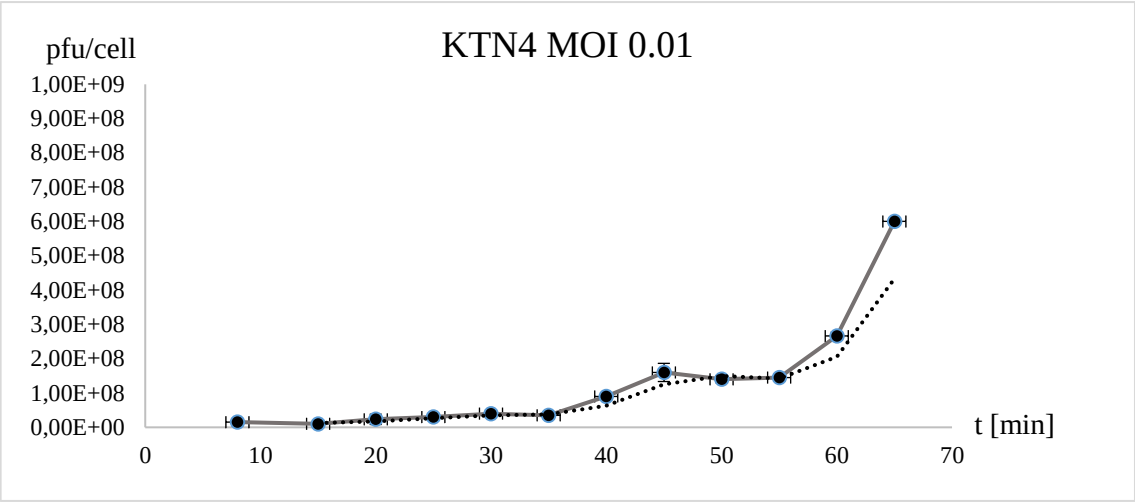


Figure 5.2: The one-step growth curve of phage KTN4. An equal volume of bacterial culture (at OD_{600} of 0.4, 10^8 cfu/ml) was mixed with phage suspension (10^6 pfu/ml) to obtain a multiplicity of infection (MOI) of 0.01. Phages were allowed to adsorb for 8 min at 37 °C . Triplicate samples were taken during 65 min at 5 min intervals and titrated. The phage titre [pfu/cell] of the solution was assessed using the double-agar layer technique, according to methods described by Adams (1959). Latent period of 40 min and burst size of 8 ± 2 phage particles per infected bacterial cell were established. A trend line is marked with dots. All data are collected from three independent experiments, standard deviations are indicated.

5.2.3. Time-lapse microscopy

In order to specifically tract in the real time the *P. aeruginosa* during KTN4 infection, a time-lapse microscopy was applied. *P. aeruginosa* PAO1 cells were grown to mid-exponential phase (OD_{600} of 0.6) and infected by CsCl purified KTN4 phage 10^{10} pfu/ml (MOI 10) on a PBS - 0.2% agar pad, for observation under the phase-contrast microscope. In the early stage of infection (between 2 and 5 min post infection) the small circular mass was observed, which migrates toward the cell midpoint and increases in size (Fig. 5.3 I, II, white arrow). It is likely the site of genome encapsidation as it was previously demonstrated by Kraemer *et al.*, 2012. After arrival at midcell, circular mass (presumably the phage nucleoid) oscillates slightly about the central axis (Fig. 5.3 IV, V). In further step of infection, bacteria cell continues to increase in size to finally achieve a significantly enlarged lemon-like shape (Fig. 5.3 V, VI). Around 30 min post infectoin, bacterial cell burst and phage progeny was released (Fig. 5.3 VI and 5, 6). The latent period of KTN4 achieved in this experiment cannot be compared to one step growth experiment, due to different conditions.

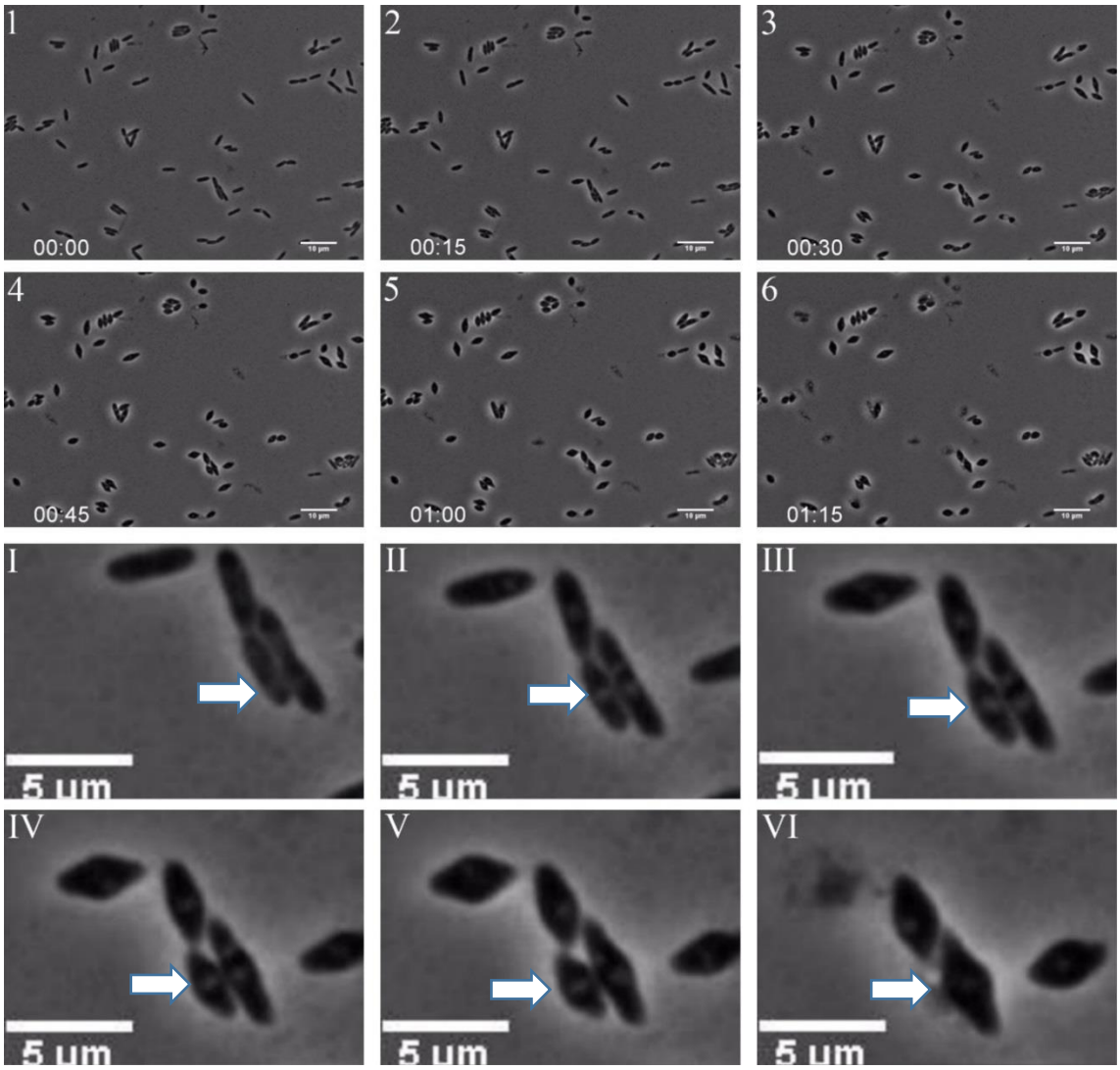


Figure 5.3: Time-lapse microscopy of the *P. aeruginosa* PAO1 infected by phage KTN4 on a 0.2% PBS-agar pad. A snapshots were taken from the start of phage infection, MOI 10, using a Nikon Eclipse Ti Time-Lapse Microscope. (1 - 6) snapshots taken every 15 min post infection, scale bare represents 10µm. (I - VI) zoom snapshots on a group of a *P. aeruginosa* PAO1 infected cells, scale bare represents 5 µm. The white arrow marks the small circular mass, that migrates from cell pole toward the cell midpoint and increases in size.

5.2.4. Stability tests

The KTN4 phage is relatively stable in a broad range of temperature and pH. No reduction in titre was observed over a period of 60 min at a temperature range of 40 - 70°C and after 30 min incubation at 80°C. After 45 min incubation at 80°C and 30 min at 90°C the phage titer decreased by 4 log units. Between 30 and 60 min incubation at 90°C the titer dropped by 6 log (Fig. 5.4). After incubation at 100°C, KTN4 particles were inactivated.

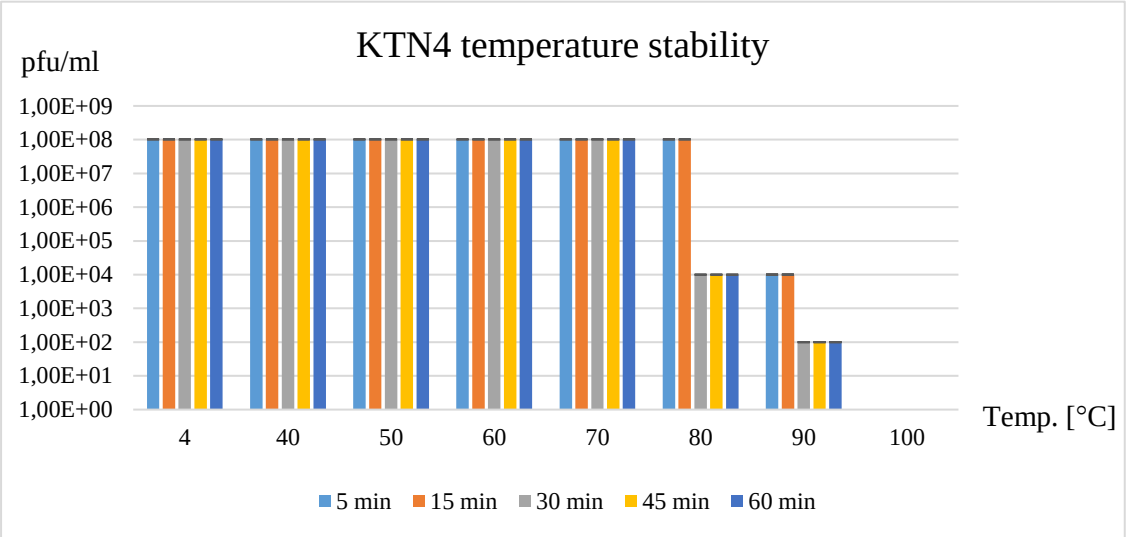


Figure 5.4: The temperature stability of phage KTN4. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

After 1 hour incubation at room temperature at pH between 6 and 12 phage remain stable. At pH 3 - 5 a 1 log and at pH 2, a 2 log reduction in phage titer was observed. At pH 1 KTN4 phage is immediately inactived (Fig. 5.5).

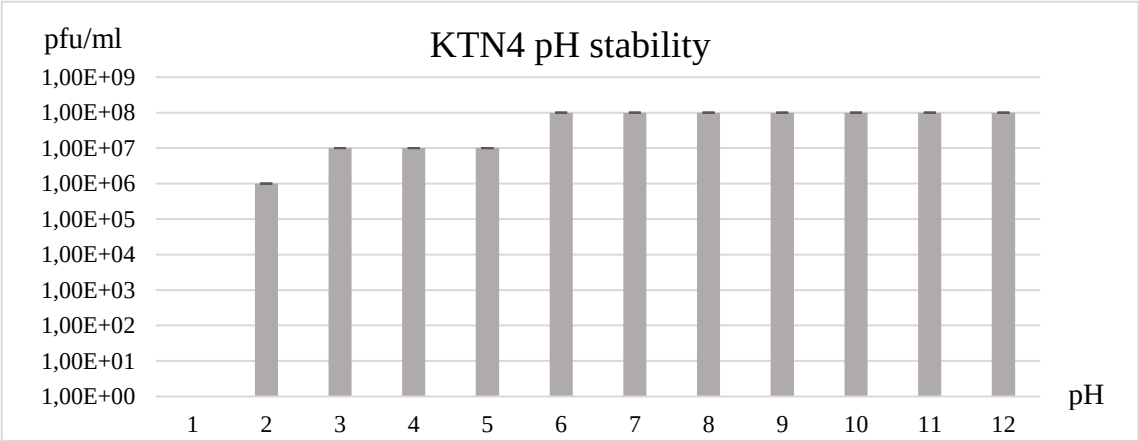


Figure 5.5: The pH stability of KTN4 phage. Phage titers (pfu/ml) were validated at the different pH values after 1h exposure at room temperature. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

Furthermore, no significant change in KTN4 titer was observed after incubation with chloroform (an equal volume of phage solution and chloroform) for 1 h incubation at room temperature and at 4 °C (Fig. 5.6).

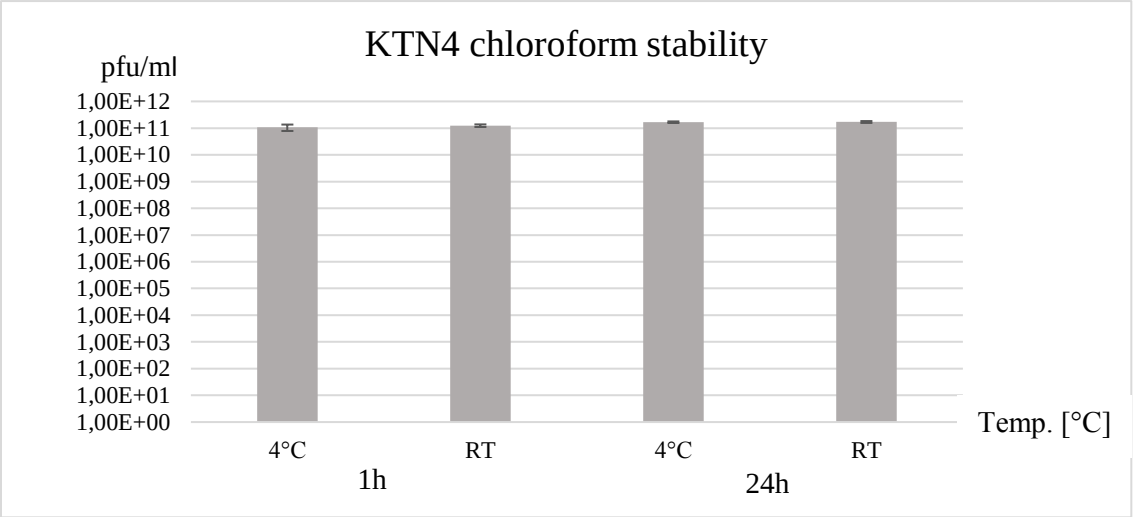


Figure 5.6: The KTN4 phage stability against chloroform treatment. Phage titers (pfu/ml) were validated after chloroform treatment for 1 h and 24 h incubation at room temperature (RT) and at 4 °C. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

5.2.5. Determination of phage receptor

In the study based on knock-out *P. aeruginosa* PAO1 mutants it was observed that KTN4 requires the IV type pili on the host cell surface as its receptor (Table 5.1).

Table 5.1. The KTN4 phage receptor identification on *P. aeruginosa* PAO1 mutants.

Bacterial strain	Phenotype	Origin	KTN4	φKZ
PAO1 (ATCC 15692)	Wild type	American Type Culture Collection	+	+
PAO1 Pirnay	Wild type with inactive type IV pili	Military Hospital Nederoverheembeek, Brussels, Belgium,	-	-
PAO1 Krylov	Wild type		+	+
PAO1 Δrmd (A-, B+)	Deficiency in D-rhamnose biosynthesis; lack of A-band LPS	Department of Molecular and Cellular Biology, University of Guelph, Canada	+	+
PAO1 ΔrmLC (A-, B-, core-)	Deficiency in L-rhamnose biosynthesis; truncate core region, lack of A-band and B-band LPS		+	+
PAO1 ΔwaaL (A-, B-)	Lack of WaaL ligating O-polymer to core-lipid A; LPS is devoid of A-band and B-band, semirough (SR-LPS, or core-plus-one O-antigen)		+	+
PAO1 ΔwbpL (A-, B-)	Lack of glucosyltransferase WbpL essential for initiation of both A-band and B-band synthesis		+	+
PAO1 ΔfliC ΔalgC ΔpilA	Lack of flagella; lack of AlgC required for A-band, core oligosaccharide, and alginate	Technical University Hamburg, Germany,	-	-
PAO1 ΔfliC wt algC ΔpilA	Lack of flagella; lack of type IV pili		-	-
PAO1 ΔfliC wt algC wt pilA	Lack of flagella		+	+
PAO1 wt fliC wt algC wt pilA	Wild type		+	+

+ activity, bacterial lysis. - no activity.

5.2.6. Host range analysis

The lytic activity of KTN4 was examined on Pirnay's *Pseudomonas* panel. This phage isolate was able to infect 35% of tested strains. However, the highly similar phage φKZ exhibits a broader host range, lysing 37% (Table S 4.1). This observation was also confirmed in the study of de Soyza *et al.* (2013), where KTN4 lytic activity was tested against the LMG *Pseudomonas* panel. In this case KTN4 was able to infect 33% isolates, compared to 46% for φKZ.

To explain host range discrepancy between these highly similar phages, several proteins involved in infection process were compared between KTN4 and φKZ genomes (a detailed description is provided in section 5.3.3.). Proteins involved in tail apparatus, lysis cassette, nucleotide metabolism and DNA replication as well as β/β′-like virion-associated non-canonical multi-subunit viral RNA polymerases (RNAP) present high similarity (99% - 100%) to their φKZ counterparts (Table 5.2, section 5.3.3.). However, when intron-encoded endonucleases and non-virion RNA polymerase subunits (nvRNAP) were examined significant variations between both phages were observed (Mesyanzhinov *et al.*, 2002, Yakunina *et al.* (2015) (Table 5.2, section 5.3.3.).

This analysis suggest, that there are no differences in host receptor recognition, destabilization of bacterial peptidoglycan as well as DNA metabolism and RNSP. Potentially the host range discrepancy could be related with variations between replication and DNA modification genes.

5.2.7. Phage influence on biofilm

To initially evaluate KTN4 biofilm eradication ability, 24 h and 48 h old *P. aeruginosa* PAO1 biofilms were grown on 8-well slide microscopy chambers and further infected with KTN4 phage suspension (8.5×10^{10} pfu/ml, 24 h incubation at 37°C). The experiment was conducted as described previously in chapter 4 section 4.2.6., where treatment with LB medium served as a negative control and a high dose of gentamicin (400 μg/ml) as a positive control. Biofilms were stained with the use of LIVE/DEAD *Bac*Light Bacterial Viability Kit and visualized by confocal laser scanning microscopy (Fig. 5.7). We observe that in both 24 h and 48 h old biofilms, KTN4 and gentamicin treatments alter the structure of the biofilm compared to the control. The untreated biofilm has a dense structure, of which the thickness increases over time. However, after both phage or gentamicin treatments, the biofilm structure is less dense

and has clear areas of biofilm disruption. This study shows the initial KTN4 antibiofilm activity, which was further evaluated with more qualitative and quantitative techniques by Drulis-Kawa group (Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland) and collaborators (Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University in Kielce, Kielce, Poland and Zeta Instruments Co., San Jose, USA). Results are published in Danis-Wlodarczyk *et al.*, 2016.

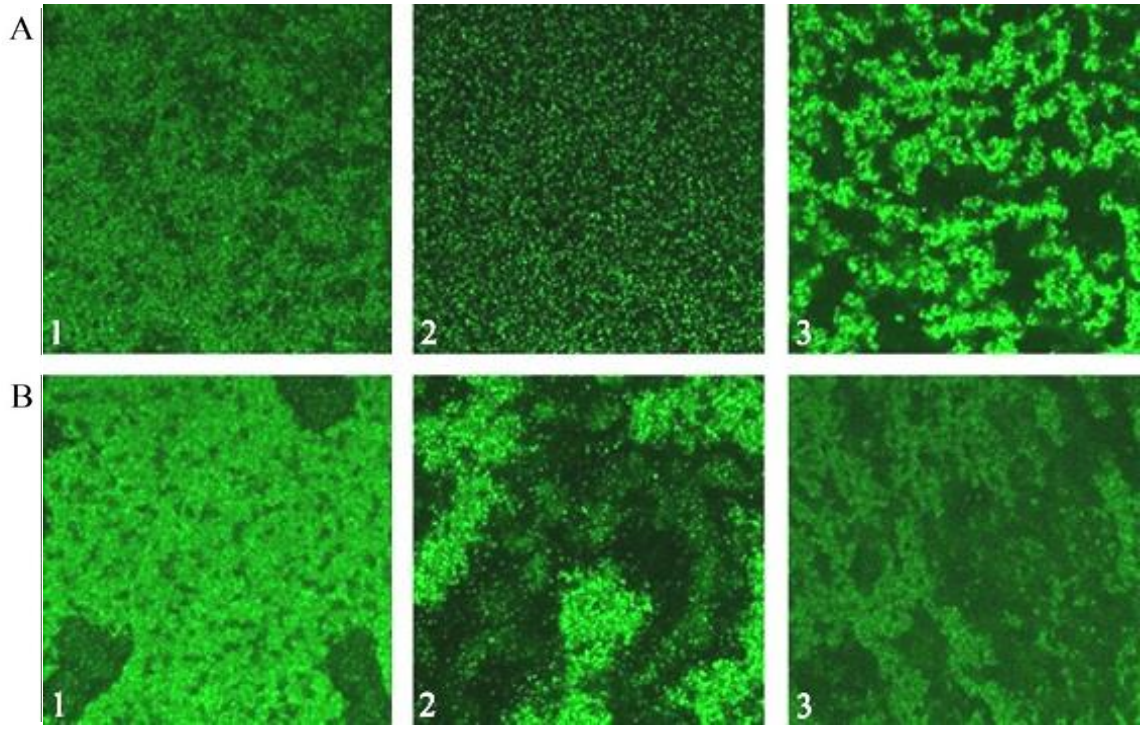


Figure 5.7: Antibiofilm activity of phage KTN4. The 24 h (A) and 48 h (B) old biofilm. 1) negative control: biofilm treated with LB broth; 2) positive control: biofilm treated with gentamicin (400 µg/ml); 3) KTN4 treatment (8.5×10^{10} pfu/ml). Biofilms were grown at 37°C, 5% CO_2, stained with the LIVE/DEAD *Bac*Light stain (LIVE/DEAD *Bac*Light Bacterial Viability Kit) and visualized by confocal laser scanning microscopy (Zeiss LSM 510 Meta 40 × objective, Jena, Germany). Live bacteria are stained green.

5.2.8. Antibacterial efficacy of KTN4 phage analyzed in gentamicin exclusion assay on Airway Surface Liquid infection model

In our study, the *in vitro* antibacterial activity of KTN4 phage was assessed in a gentamicin exclusion assay on Airway Surface Liquid (ASL), which is to our knowledge, the first report showing phage treatment efficacy in that infection model. For experiments, two cell lines were selected: 1) NuLi-1 derived from normal human bronchial epithelium and 2) CuFi-1 derived from CF patient bronchial epithelium with

significantly thinner ASL (Zabner *et al.*, 2003). Three *P. aeruginosa* strains were selected for these experiments: PAO1 reference strain (piliated, motile strain, effective biofilm former), a nonCF0038 isolate from burn wound (highly expressing type IV pili), both reflecting CF early colonizing isolates, which are non-mucoid with typically smooth LPS and more virulent. The third strain was CF708 from late infection phase, presenting slowly growing and less virulent small colony variants with low expression of type IV pili, and biofilm formation (Olszak *et al.*, 2015).

Several controls of epithelial cells viability were prepared at different points of the experiment: (i) a negative control without any treatment, (ii) a negative control with TC media; (iii) a positive control with Triton-X100; (iv) 1.5 h after KTN4 application; (v) 1.5 h after *Pseudomonas* infection. The Nuli-1 and CuFi-1 cells were stained with 8 μM Calcien AM (live staining) (Life Technologies, NY, USA) and 3 μM propidium iodide (PI) (dead staining) (Life Technologies, NY, USA) according to manufacturer's instructions. After staining, in all cases, filter inserts were XZ scanned using a confocal microscope (Zeiss LSM 510 Meta 40 × objective, Jena, Germany). No toxicity influence of the cell lines was observed for phage and bacteria samples after 1.5 h (Fig. 5.8).

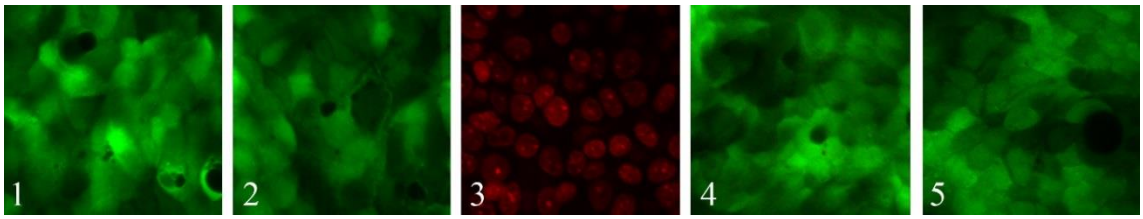


Figure 5.8: Live/dead staining of lung epithelia cells prepared at different time points of experiment. (1) a negative control without any treatment, (2) a negative control with TC media; (3) a positive control with Triton-X100; (4) 1.5 h after KTN4 application; (5) 1.5 h after *Peudomonas* infection. Cells were imaged using confocal microscopy (Zeiss LSM 510 Meta 40 × objective, Jena, Germany).

First, both epithelial cell lines were infected with selected strains for 3 h and colony count showed *P. aeruginosa* efficiently propagated in both ASLs (10^7 - 10^9 cfu/ml) (Fig. 5.9). Although PAO1 and nonCF0038 pathogens grew better on NuLi-1 cell line, reaching around 1 log higher compared to CuFi-1 ASL, no statistically significant differences have been found ($p > 0.05$). The small colony variant CF708 grew equally well in both types of mucus layer. In the second step of experiment, the KTN4 phage eradication ability of the extracellular bacterial load was evaluated. The CFU counts of *P. aeruginosa* were significantly ($p < 0.05$) reduced for normal NuLi-1 epithelia cells. A 7 log, 6 log and 4 log decrease was observed for PAO1, nonCF0038 and CF708, respectively (Fig. 5.9 A). In the case of the CuFi-1 epithelia, the phage treatment was also

very effective, giving 4 log, 6 log and 5 log reductions in the colony count of PAO1, nonCF0038 and CF708 ($p < 0.05$), respectively (Fig. 5.9 A). PAO1 was significantly more susceptible ($p < 0.05$) to phage treatment in NuLi-1 cells compared to CuFi-1 cells, contrary to the CF708 isolate. It was confirmed that the KTN4 phage could freely diffuse and gain access to the bacterial hosts in both ASL models, but the outcome of the treatment was strongly dependent on the strain features or were influenced by the differences in ASL pH between NuLi-1 and CuFi-1 cells. A possible explanation of this phenomenon was observed by Worlitzsch's studies (2002) where *P. aeruginosa* was not interacting with the CF epithelium directly, but was rather found trapped in mucus plugs formed in the airways. Thus, the phage receptors could be masked by mucus elements, which have an influence on phage adsorption to bacterial cell surface. Moreover, the CF strain is better adapted to CuFi-1 environment and may presumably express the type IV pili more efficiently, which are receptors for KTN4 phage.

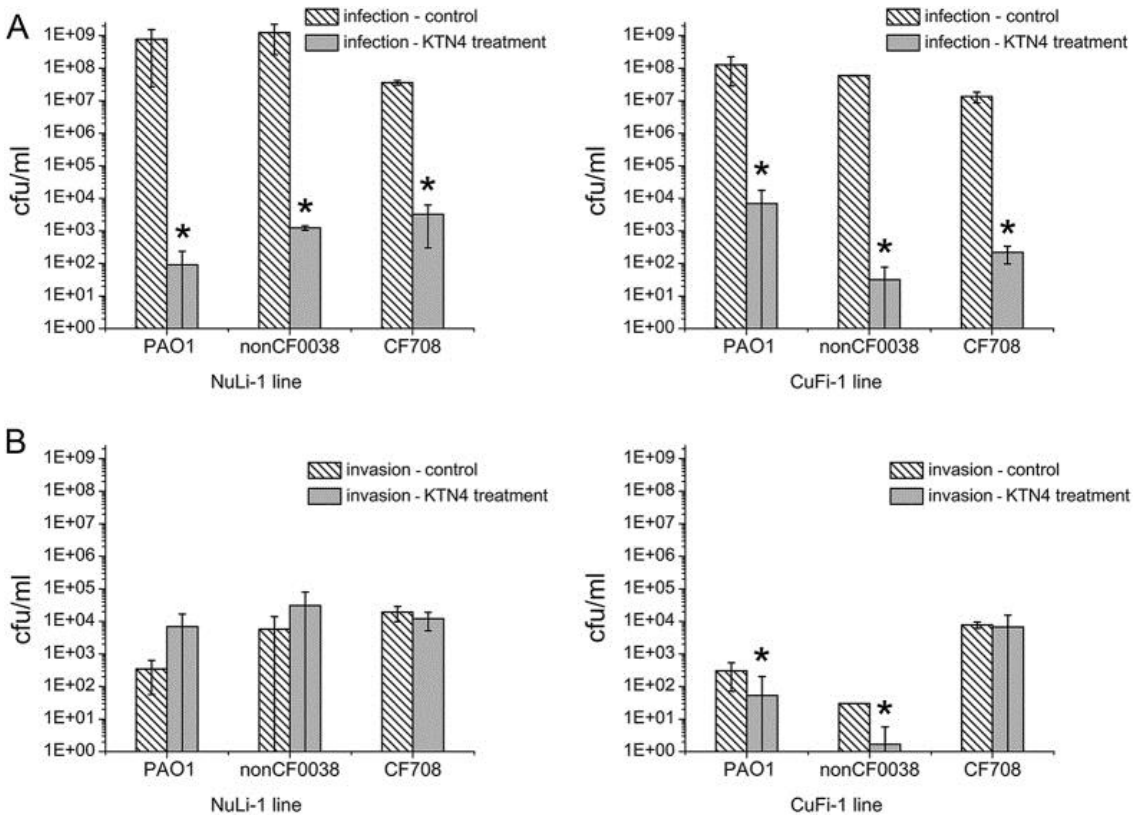


Figure 5.9: Phage KTN4 treatment of *P. aeruginosa* infected NuLi-1 and CuFi-1 epithelial cells. (**A**) Colony count of bacteria collected from apical wash; (**B**) colony count of bacteria internalized in epithelial cells. The results are presented as the means ± SD. Statistical analysis was made by the ANOVA test (denoted *p*-values < 0.05) (Danis-Wlodarczyk *et al.*, 2016).

In the next step of our experiment, the ability of *P. aeruginosa* strains to invade into epithelial cells was investigated (Fig. 5.9 B). Although *P. aeruginosa* was generally thought to be an extracellular pathogen, a number of different groups have found that

it can be internalized into a range of different cell types, including epithelial cells (Schroeder *et al*., 2001; Fleiszig *et al*., 1995). The results showed that CF and non-CF strains could indeed be internalized into both cell lines. No statistically significant differences ($p > 0.05$) were observed for all pathogens. Nonetheless, the PAO1 and nonCF0038 were less internalized into the epithelial cells compared to the CF708 isolate, where only 4.50E-05% and 2.00E-04–5.00E-05% of infecting population were able to internalize into NuLi-1 and CuFi-1 cells, respectively, in contrast to 0.05% and 0.06% for the CF708 isolate. Therefore, internalization capability and adaptations of CF708 strain was more effective compared to the nonCF strains, consistent with previous observations (Fleiszig *et al*., 1995). This suggests that the *Pseudomonas* strains with high cytotoxicity are low invasive and vice versa, bacteria that are less virulent enter the epithelial cells to survive intracellularly without killing the host cell.

In the final step, the influence of KTN4 phage treatment on the number of invaded bacteria was evaluated (Fig. 5.9 B). In the NuLi-1 cell line there were no significant changes ($p > 0.05$) in CFU counts for all *P. aeruginosa* invading strains after phage application. A possible explanation is that during the 3 h of pretreatment, all bacterial cells were already internalized, thus the phage had no access to these host cells. In contrast, the phage application was significantly more effective ($p < 0.05$) for CuFi-1 internalization prevention by wild type *P. aeruginosa* strains, since CF lung cells due to lack of CFTR, show significantly less ingestion rate of LPS-smooth bacteria and significantly greater lung burdens post-infection than wild-type epithelium (Schroeder *et al.*, 2001). As previously mentioned (Worlitzsch *et al*., 2002), wild strains not well adopted to the ‘sticky’ and dense CF environment are absent on epithelial surface. They remain as macrocolonies within intraluminal material, slowing down efficient internalization. Simultaneously, phage application cause effective eradication of bacterial cells trapped in the mucus plugs. The CF708 isolate evolving in CF lung environment is able to internalize with the CuFi-1 epithelium in relatively short time after infection evading phage lytic activity.

5.3. Genome analysis

5.3.1. Genome properties

Using high throughput sequencing by the Illumina MiSeq platform, the complete genome sequence was determined. KTN4 has a linear, circularly permutated and terminally redundant, A+T-rich (36.9% GC) double-stranded DNA molecule (279,593 bp) (Fig. 5.10).

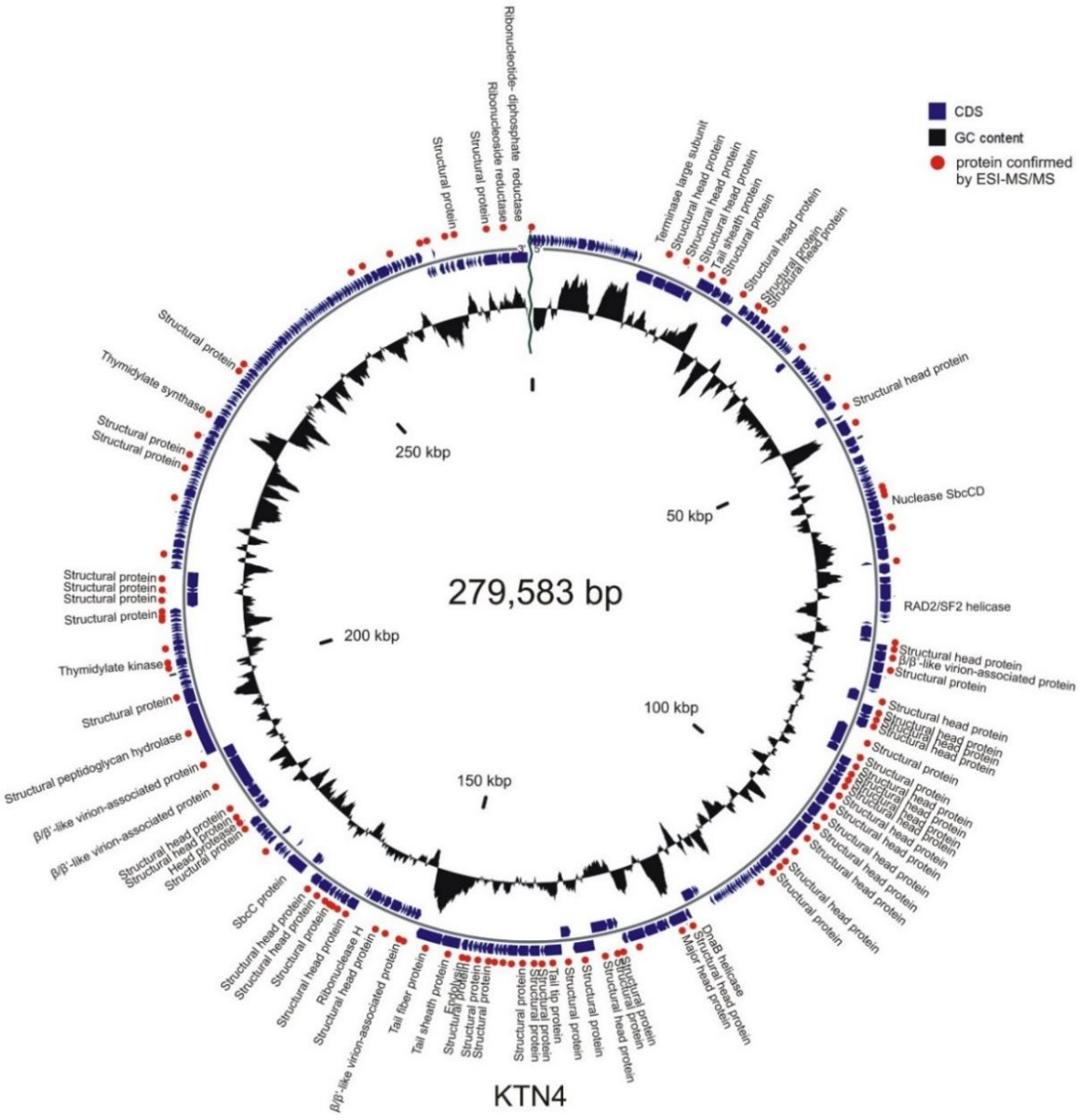


Fig. 5.10: Circular representation of the KTN4 genome. Blue arrows represent genes with predicted functions, grey arrows represent hypothetical proteins. The inner circle represents the GC content. Structural proteins, which presents was confirmed by ESI-MS/MS analysis are marked with red dots. The KTN4 genome was deposited at GenBank under accession number KU521356.

In total, 368 ORFs could be predicted, varying in size from 36 to 2237 amino acid residues (Table S 5.1), as well as six tRNAs (Leu (UUA), Pro (UUG), Met (CAU),

Asp (GUC), Asn (GUU), Thr (UGU). Only 87 proteins have a predicted function. According to the orientation of transcription, the ORFs are organized into 134 operons most of which are on a positive strand. The KTN4 shows a genome-wide nucleotide sequence similarity to: φKZ 99%, PA7 99%, phiPA3 84%, 201phi2-1 78% (BLASTN). As such, it can be defined as an isolate of the *Pseudomonas* phage φKZ species. However, there are a few significant differences. Nineteen additional proteins of unknown function are present in KTN4 genome (gp14, gp23, gp24, gp25, gp26, gp30, gp31, gp32, gp34, gp59, gp75, gp93, gp94, gp150, gp286, gp287, gp315, gp321, gp357) and absent in the φKZ genome. Also 17 φKZ proteins of unknown function are missing in KTN4 genome (gp24.1, gp24.2, gp70.1, gp76.2, gp117.1, gp117.2, gp144.1, gp179, gp232.1, gp239, gp260.1, gp290.1, gp294.1, gp294.2, gp295, gp295.1, gp296).

Based on φKZ RNA-seq analysis performed previously by Ceyssens *et al.* (2014) and using the PISE EMBOSS fuzznuc program and MEME-MAST, 47 promoters were predicted for phage KTN4. Among them, 31 are early phage-specific promoters with highly conserved, uni-directionally distributed AT-rich intergenic motifs (5'-TATATTAC-3') (Fig. 5.11). Furthermore, less conserved upstream (5'-TTTaA-3') and downstream (5'-TG-3') motifs were found. The middle promoters are located on both strands and distributed throughout the whole genome. They are linked by only a weak AT-rich motif (5'-AAanntTAC-3'; lowercase letters represent a lower level of conservation) centered at position 24 with respect to the transcription start site (Fig. 5.11 B). For late transcription no sequence conservation upstream of 5' ends could be detected apart from a 5'-TATG-3' motif overlapping the transcription start site (11 late promoters) (Fig. 5.11 C). This findings are consistent with previous studies of φKZ phage (Ceyssens *et al.*, 2014).

Using ARNold, 107 of putative factor-independent terminators were predicted. Most potential stem loop transcription terminators contain the tetranucleotide UUCG loops (Fig. 5.12).

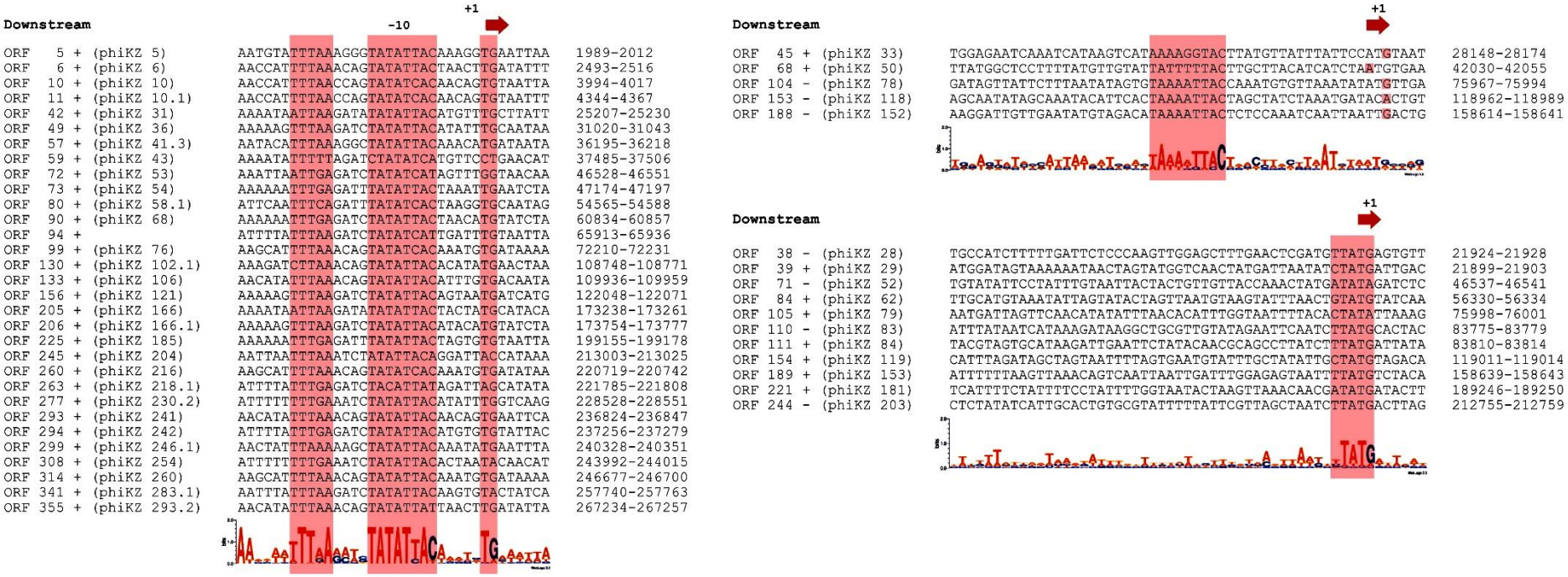


Figure 5.11: Alignments of KTN4 promoters. (A) Early promoters. The 5’ ends of primer extension products, which correspond to the transcription start sites, are located ~ 10 bp downstream from the center of the core TATATTAC motif and are associated with an additional conserved 5’-TG-3' motif. (B) Middle promoters. They are united by only a weak AT-rich motif (5’-AAanntTAC-3’; lowercase letters represent a lower level of conservation) centered at position 24 with respect to the transcription start site (C) Late promoters. No sequence conservation upstream of 5’ ends of late transcripts could be detected apart from a 5’-TATG-3’ motif overlapping the transcription start site. The corresponding sequence logos are depicted below the alignments. Pink bars delineate conserved promoter elements.

KTN4 Terminators

ORF	4	+	Erpin +	ACATAATAATAGCCTTCCCCCAGGGGAAGGCtTTATGTCATGTAT	-17.40 1946-1965
ORF	8	+	Both +	TGGCGACATACACCCTCCCTTCGGGGAGGGCTTTATTTTCAATA	-17.80 3569-3590
ORF	16	+	Both +	CACGCAATAATGGGAGTCTTCGGACTCCCTTTTATTTTCTA	-16.10 7171-7188
ORF	18	+	Both +	AACGATATAAATACCCTCCTTCGGGAGGGTATTTATTTTGTGA	-17.50 8763-8782
ORF	33	+	Rnamotif +	GGCTATATGTCGTCAAGTTTAAGCTTGACTTTTTGCCCAAA	- 8.70 13835-13855
ORF	35	-	Both -	TTAAATATAAAGCCCCCTAGGAACCACCTAGGGGGCTTTATAACGTCTT	-16.20 16500-16476
ORF	41	+	Both +	TAAAGCATATCGCCTCCCTTCGGGGAGGCTTTATGTTGTTA	-18.50 25158-25175
ORF	43	+	Both +	TAACAGCATAAGCCCTCCCATTGCGGGAGGGCaTTATGTTTATTT	-18.90 26845-26865
ORF	44	-	Both -	ATAAACATAATGCCCTCCCGCAATGGGAGGGCTTATGCTGTTAAG	-17.80 26865-26845
ORF	46	+	Erpin +	GTAAATATAATGCCCTCTCCATAAGGAGAGGGCTTATGATGTATTT	-17.70 29290-29311
ORF	48	+	Erpin +	ATGAACATATTGCCTCCCCTAGGGGAGGCTTTATGCTGAGTT	-16.30 30981-30998
ORF	52	+	Both +	TAACTTATAATGGAGCCCTTCGGGGCTCCTTTTATTTTTTT	-18.20 34542-34559
ORF	55	+	Both +	TCGAAAGAGTCCACCCTGGATGGGGTGTTATATTAACTG	- 9.00 35551-35566
ORF	57	+	Both +	TAATGGCATAAGCCCTCCCATAATGGGAGGGCaTTATGTCTTTTT	-19.20 36537-36557
ORF	58	-	Both -	AAAGACATAATGCCCTCCCATTATGGGAGGGCTTATGCCATTAAG	-18.50 36557-36537
ORF	60	+	Erpin +	TACTAAATAAAGCCCTCCTTATGGGGGGCTTTATACCGATTA	-14.70 39035-39054
ORF	67	+	Both +	ATAACTATAAAGGAGCCATTATGGCTCCTTTTATGTTGTA	-11.80 42000-42016
ORF	68	+	Erpin +	TAAGAAATATAGCCTCCCCTAGTGGGAGGCTTATATACTAAAA	-14.40 44397-44415
ORF	70	+	Both +	ACGGCATATAACCCCCTCCAATTTAGGAGGGGGTTTATGTCACATT	-18.20 45385-45406
ORF	71	-	Erpin -	TGTGACATAAACCCCCTCCTAAATTGGAGGGGGTTATATGCCGTTT	-18.70 45406-45385
ORF	73	+	Both +	AAACTGATATAGCGGCTCCTTCGGGGGCCGCTTATATCTCCTT	-20.00 49476-49495
ORF	77	+	Both +	AAATAAATTAATAGGGGAGTTAGACTCCCCTATTATTTCTTTGA	-14.50 53627-53646
ORF	81	+	Erpin +	TTAAATAAATGAGAGTCCCGTACAGGGGCTCTcTTCTATCGTTTAG	-13.80 55484-55504
ORF	92	+	Both +	CACGGCATAAACCAGGGGGCCTAAGGGTCCCCTGGcTTGCTTTAAAAC	-15.70 65499-65522
ORF	98	+	Both +	GGCATAAATGATCCCTCCCTTAGTTGGGAGGGATTATGCTTTTTT	-17.10 72155-72176
ORF	101	+	Rnamotif +	GGCGACATAAACCCCTCTCCATGTGGAGAGGGGTCTGTCTTTATT	-17.50 73283-73304
ORF	104	-	Rnamotif -	AATAATGACGAGGTGGAAACTCCACCTTTCGTTCCTGA	- 8.50 75265-75251
ORF	107	+	Erpin +	GTAAACATATTGCCTCCCCTAGGGGAGGCTTTATGCTGATTA	-16.30 79747-79764
ORF	109	+	Both +	CACATCATAAAACCCTCCCATTACGGGAGGGTTTATTTCGTTAAC	-15.70 82201-82220
ORF	110	-	Both -	TAACGAAATAAACCCTCCCGTAATGGGAGGGTTTTATGATGTGTC	-15.30 82220-82201
ORF	113	+	Both +	TACAACATAATGCCCTCCCCTAGGGGAGGGCTTATGACCTTATT	-19.60 86840-86859
ORF	120	+	Both +	ACAAACATAATGAGGAACCCTTCGGGGTTCCTCTTATGCTATGTAA	-15.30 96180-96200
ORF	122	+	Both +	GACGACATATTGACGGGGCTCGGAAGAGTCCCGTCTTATGTTAGGAG	-14.90 99507-99530
ORF	124	+	Both +	ACAAACATAACGGGGAGCCTTCATGGCTCCCTTTTATGCTGTTAG	-16.20 103225-103244
ORF	136	+	Both +	TAACTGATAATGGGAGTCTTCGGACTCCCTTTTATTTTCTA	-16.10 111420-111437
ORF	142	+	Erpin +	ATAGAAAAAAAACCACCCTATATAAGGGTGGTTTATAAAAGAGGT	-13.30 113888-113908
ORF	150	+	Rnamotif +	TGAATCGGATCGCCCCGTAGGGGCTTTTATTTCATA	- 9.10 116426-116438
ORF	151_1	-	Rnamotif -	AATCTACACTGTGTCTCCTTTTAGGGAGACAgTATGTATTAGCT	-12.10 116550-116531
ORF	151_2	-	Rnamotif -	AAAGCATTTAAGCCCTCATAGGAGGGCTTTTATGTCGTA	-11.70 116704-116689
ORF	155	+	Both +	AACAGACATAGGCCCTCCCTTCGGGGAGGGCTTTATGTTGTAT	-21.50 122002-122021
ORF	159	+	Both +	TAAGCATGAAAGCCCCTCTTCGGAGGGGCTTATTTATGAGT	-18.40 127450-127467
ORF	160	-	Both -	TAAAGACATAGCCCCTCTCCATGCGGAGAGGGGTTTATGCCATTAC	-17.20 128196-128175
ORF	164	+	Erpin +	AACAACATAATGCCCTCCCATTACGGGAGGGCTTATGACGTTAAC	-18.20 134466-134486
ORF	165	-	Both -	TAACGTCATAAGCCCTCCCGTAATGGGAGGGCaTTATGTTGTTTAG	-18.50 134486-134466
ORF	167	+	Both +	ACCCGGCATAAGAGACACAGCCAAAAGCTGTGTCTCTTTTTGCTGTTT	-11.70 138555-138579
ORF	178	+	Both +	AAAACTATAAAGCCTTCCTCTTCAAGAGGAAGGCTTTATAATGCTAT	-14.20 147015-147037
ORF	180	+	Rnamotif +	TGTAACTTTTTGAGAGTCCTTCGGGACTCTCTTATGTTGTTAA	-17.30 148547-148566
ORF	182	+	Both +	AACGACATATTGCCCTCCCTTCGGGGAGGGCTTTATTTTGTCT	-21.50 154331-154350
ORF	189	+	Both +	TAACAGCATAAGCCCTCTCCCAAAGGAGAGGGCaTTATGTCGTACCA	-18.40 159606-159627
ORF	190	-	Erpin -	TACGACATAATGCCCTCTCCTTTGGGAGAGGGCTTATGCTGTTATC	-17.70 159627-159606
ORF	194	+	Both +	TGCATAAATGAGAGAGGGCAAAGGCCCTCTCTTATATTAATCT	-16.20 163817-163836
ORF	197	+	Both +	TTTTAATTTAAAGCGGATCTGCTTTAAAGTGGATCCGCTaTTATTTCGTCTTA	-12.70 165695-165722
ORF	200	+	Both +	TAACAGCATAAGCCCTCCCATAGTGGGAGGGCaTTATGTTCTTTA	-19.20 169233-169253
ORF	201	-	Both -	AAGAACATAATGCCCTCCCACTATGGGAGGGCTTATGCTGTTAAA	-18.50 169264-169530
ORF	203	+	Both +	AAGCAACATAAGCCCTCCCGCAATGGGAGGGCaTTATATTAGTAA	-18.50 172994-173014
ORF	204	-	Both -	ACTAATATAATGCCCTCCCATTGCGGGAGGGCTTATGTTGCTTAT	-18.20 173014-172994
ORF	209	-	Erpin -	AACGACATATAGCCCTCCCGTAATGGGAGGGCgTTATGCACTTAAT	-18.50 175215-175195
ORF	214	+	Both +	TGACAGCATAAGCCCTCCCATTACGGGAGGGCaTTATGTCTTTTT	-18.90 179703-179723
ORF	215	-	Both -	AAAGACATAATGCCCTCCCGTAATGGGAGGGCTTATGCTGTCATG	-17.80 179723-179703
ORF	218	-	Rnamotif -	GGCACCACCAGCTCCCCATCTCTGAATGGGGAGTTTTTAACTGAG	-13.20 183101-183080
ORF	221	+	Both +	AATCTATTAAAGGAGCCTCCCAATAGGGAGGTTCCTTATTTTGTCAA	-15.30 196004-196027
ORF	224	+	Both +	TTGACATAAATGCCTCCCTTCGGGGAGGCTTTATTTCGTAT	-18.20 199116-199133
ORF	227	+	Both +	CAATATAATAAGATCAGTGGGTGCATTGCCCACTGGTCaTTTTTTATATTAA	-12.00 200478-200504
ORF	230	+	Both +	ACGATATAATAGAGTCCCCATTACGGGGACTCTTCTTTATAATT	-16.30 202265-202285
ORF	238	+	Both +	GTTATATATTACCCTCTCCTTCGGGAGAGGGTTTTATGTTTAA	-19.20 206302-206321
ORF	239	+	Rnamotif +	TACTTTTATAGGAGAGAACCATGATGGTTCTCTCTTTTTGACGTAA	-12.70 206931-206953
ORF	241	+	Both +	AACAACATAATGCCCTCCCATTGTGGGAGGGCTTATGCCGTTAAA	-18.50 208132-208152
ORF	242	-	Both -	TAACGGCATAAGCCCTCCCACAATGGGAGGGCaTTATGTTGTTTTA	-19.20 208152-208132
ORF	245	-	Both -	AATGGCATAAAGGAGCCCCTGGTGGGCTCCTTTAATGTAATAG	-14.40 212979-212959
ORF	247	+	Erpin +	GATAACATATACCCTCCCCTTTATTGGGGAGGGaTTATTTACTTATA	-19.20 214043-214064
ORF	249	+	Erpin +	AGCCAAAATAAACCCTCCTCTTAGTAGGAGGGTTTATATCGTCTCA	-13.20 215669-215690
ORF	261	+	Erpin +	TGTTGTCATAAAGCCCTCCCTTCGGGGAGGGCTaTTGCACTAAATTT	-17.90 221747-221768
ORF	263	+	Both +	TTTGAACAATAAGCGGTGTGTACACCGCTtTTATTCCGTGTTT	-11.70 222145-222162
ORF	265	+	Rnamotif +	TTCGAGGCATATACTCCATGACTTATGTTGTGGGGTATATGTGTTATAA	- 8.50 223844-223869
ORF	267	+	Both +	GTTTAATATTAAGGGACTCTTAATGAGAGTCCCTaTTTATTTTGCGTA	-14.10 224652-224673
ORF	269	+	Erpin +	AAGACTGGGCAAAAGACATTCTAAAAGAGTGTTGGCTGCCCAGTCaTCTTTTCTGCCGT	-14.93 225691-225733
ORF	274	+	Both +	GTAAAATAAATACCAGCCCTTCGGGGCTGGTTTATTTCATCTAG	-18.70 227846-227865
ORF	276	+	Both +	AAGCATAAATAGGAGAGCTATTTAGCTCTCCTTTATTGCATTAA	-14.10 228489-228508
ORF	284	+	Both +	TAACATAAAACGCCTCCCTTCGGGGAGGCTTTATTTTCCTT	-18.50 233224-233241
ORF	285	+	Both +	AACAGTATATCGGGAGTCCATTTGGACTCCCTTTTATATCGTA	-16.70 234194-234213
ORF	291	+	Both +	GATTGCATAAAGAGAGAGGATAAATTCCTCTCTCTTTTTATGTTGT	-14.40 236116-236138
ORF	292	+	Both +	TACCATAAATACCTCCCTTCGGGGAGGTTTTATATTGTG	-14.70 236776-236791
ORF	293	+	Both +	TAAGAGCATAAAGCCCTCCCTTCGGGGAGGGCTaTTGCACTAAATTT	-17.90 237218-237239
ORF	296	+	Both +	TTGTATAATAAGAGCCCTTCGGGGCTCTTTCTATTTCGC	-15.10 239090-239105
ORF	298	+	Both +	ACTTAATATACACCCTCCCTTCGGGGAGGGTtTTATTTATCTTTT	-19.00 240277-240296

ORF 306 +	Both +	AGAAATAATTAGAGGACCTTCGGGTCCTCTTTTTAACCGAC	-16.30	243758-243775
ORF 307 +	Both +	AAAGTCATAAACCCTCTCCCAAAGGAGAGGGcTTTTACATTAAA	-15.20	243954-243974
ORF 313 +	Both +	GTTATTATATTGCCCTCCCATTGCGGGAGGGCTTATATTTTATT	-18.20	246626-246646
ORF 319 +	Both +	ATTATGAAAATGGGGGTCTAAAGACCCCCTATCTATTTAAT	-15.10	249167-249184
ORF 322 +	Both +	ACGCCATAATGGCCTCCCTTCGGGGAGGCTTTTATATCCCT	-18.20	250436-250453
ORF 329 +	Both +	ATGCTTAATTAGGGAGGGTAATACCTCCCTTTTTTATATAA	-12.80	252726-252743
ORF 334 +	Both +	ATGGTATAATTGCCCTCCTTCGGGAGGGCTTCTATTTCGTA	-18.20	254927-254944
ORF 342 +	Both +	AATCTAATACATAAAGGGCTCTTCGGAGCCCTTTATTTTTCTTTAT	-11.70	259065-259088
ORF 347 +	Both +	CCAATGAGATACCCTCCCCTGATCGGGGAGGGTTTTTGACATCT	-17.90	263102-263122
ORF 348 +	Both +	CTATTGATAATGCCCTCCCTTCGGGGAGGGCaTTATTTCGCTTAA	-22.00	263633-263652
ORF 351 +	Rnamotif +	TACGGCATAAACCCCTCTCCGTATGGAGAGGGGaTATGTCTTTATT	-18.00	265810-265831
ORF 352 -	Both -	TAAAGACATATCCCCTCTCCATACGGAGAGGGGTTTATGCCGTAAT	-17.10	265831-268510
ORF 355 +	Both +	AAGATAACATAGAGCTAGCCCATTAGGGGCTAGCTCTTATATTTATTT	-14.70	267454-267478
ORF 356 -	Rnamotif -	AATAAATATAAGAGCTAGCCCCTAATGGGCTAGCTCTATGTTATCTTA	-14.70	267478-267454
ORF 357 -	Both -	AGCTAATTATAACCCTCCCATTATGGGAGGGTtTTTATCAGGTTAT	-16.00	268288-268268
ORF 358 -	Both -	AATAAAATAAAAGCCCTCCATCACGGAGGGCTTTATGTTGTTATG	-15.70	268954-268934
ORF 359 -	Rnamotif -	TGTGTCTGAATGCCAGGGGTTTGCCACCCCTGGCTATATTATGTCA	-18.00	269672-269650
ORF 360 -	Both -	CATATTCTAAAGAGTAGCTTCGGCTACTCTTTTATGTTGCCA	-14.10	270872-270855
ORF 363 -	Both -	CCATATACTAGCCTCCCACTTGGGGGAGGTTTTATACTGTC	-11.20	272004-271987
ORF 365 -	Both -	GACGATATATCGCCTCCCTTCGGGGAGGCTTTATATTTTTTT	-18.50	273178-273161
ORF 366 -	Erpin -	TAGACATAATGACCCTCCCTTAATTGGGAGGGTTTATGCTAACAAT	-16.50	275278-275257
ORF 368 -	Both -	TTAAGTGTATACACCCTCCCTTAATTGGGAGGGTGTTATTCCCAATTT	-17.10	276961-279249

Figure 5.12: Phage KTN4 predicted terminators with palindromes marked blue. From the left site: list of ORFs after which terminators are located, strand and motif of terminator, terminator sequence with conserved stem-loop structures marked blue, ΔG (the Gibbs free energy of stem-loop formation in kcal/mole) and location in phage genome.

5.3.2. ESI-MS/MS proteome analysis

Using ESI-MS/MS analysis of proteins from denaturated phage particles fractionated on SDS-PAGE, 111 gene products have been identified, among which five virion-unrelated enzymes, 36 virion associated proteins and 70 structural gene products, with sequence coverages ranging between 5.9% to 89.0% (Fig. 5.13, Table S 5.2).

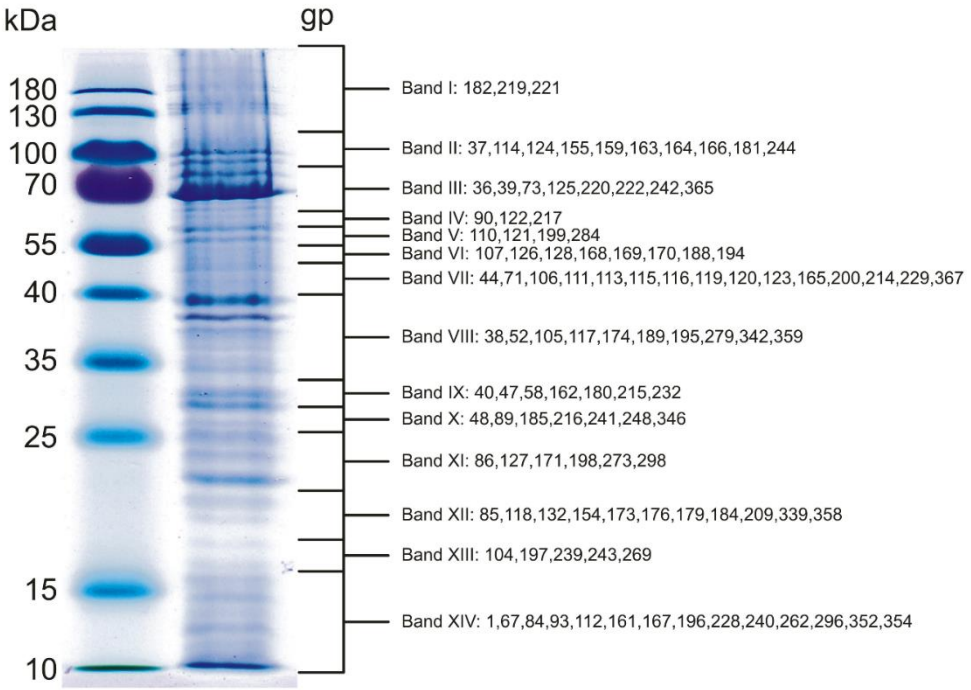


Figure 5.13: The SDS-PAGE pattern of KTN4 structural proteome against PageRuler Prestained Protein Ladder (Thermo Scientific) in the first line. The corresponding molecular weight is mentioned left. The numbered fractions on the right, correspond to gel slices analyzed individually by ESI-MS/MS. The proteins are mentioned in the slice in which they were most abundantly present. The 12% SDS-PAGE gel stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific).

The KTN4 structural proteins were compared to their homologues from φKZ (NC_004629.1 from 2008 and AF399011.1 from 2013), PA7 (JX233784.1), phiPAK3 (HQ630627.1) and 201phi2-1 (NC_010821.1) phages (BLASTP). As expected, the highest similarity was found for φKZ structural proteins, ranging from 63% to 100%, except for gp93, which was identical to PA7 hypothetical protein (AFO71119.1).

5.3.3. Functional annotation of KTN4 genome

Capsid

ESI-MS/MS analysis identified 31 structural head proteins (97 - 100% similarity with φKZ counterparts) including gp215 head protease identical to φKZ gp175 and a major head protein gp155 with 98% similarity to φKZ gp120 (Thomas *et al.*, 2012) (Table S 5.2).

Tail apparatus

The contractile tail of φKZ, the closest homologue of KTN4 (98 - 100% similarity), is built from at least 32 different proteins, but a definitive structural function was assigned to only two of them: the tail sheath protein (KTN4 gp39 identical to φKZ gp29, ESI-MS/MS hit) and the tail tip protein (KTN4 gp166, 99% similar to φKZ gp131, ESI-MS/MS hit) (Table 5.2). The φKZ gp131, located at the periphery of the baseplate and possibly associates with fibers that emanate from baseplate (Sycheva *et al.*, 2012), has several paralogs in the φKZ genome: gp132, gp133, gp134, and gp135 (KTN4 gp167, gp168, gp169 and gp179, respectively, 99% similarity). They show a notable sequence similarity at their N termini and appear to form a paralogous cluster, a common characteristic of tail fiber/tail spike genes. Sycheva *et al.* (2012) hypothesized that gp131 might play a role in host cell recognition by binding to some cell surface receptors, however does not participate in irreversible adsorption of the φKZ particle to PAO1 cells in the laboratory conditions, neither cleaves the peptidoglycan nor the expolysaccharide of PAO1 cells. Moreover, in the ESI-MS/MS analysis two additional proteins were identified gp181 (putative tail sheath protein, φKZ gp145, 99% similarity) and gp182 (putative tail fiber, φKZ gp146, 98% similarity).

Lysis cassette

The KTN4 endolysin (gp180) is highly identical (99%) to the endolysin of *P. aeruginosa* φKZ gp144. This enzyme is well studied both at the molecular,

biochemical (Briers *et al.*, 2007, 2009; Paradis-Bleau *et al.*, 2007; Cloutier *et al.*, 2010) and structural level (Fokine *et al.*, 2008). It comprises a N-terminal peptidoglycan binding domain that has high binding affinity for its substrate (295 nM) (Briers *et al.*, 2009) and a C-terminal transglycosylases domain that cleaves the glycosidic linkage between N-acetylmuramic acid and N-acetylglucosamine with the concomitant formation of 1,6-anhydromuramic acid (Paradis-Bleau *et al.*, 2007). Moreover, the ESI-MS/MS analysis identified the presence of tail associated enzyme gp221 (structural peptidoglycan hydrolase), that correspond to gp181 of φKZ phage (99% similarity) (Table 5.2). This enzyme cleaves the host cell wall during the first stage of the life cycle (Briers *et al.* 2008). In the centre of the φKZ baseplate, there is a density that resembles the needle-like ‘‘cell-puncturing’’ device of T4, which is most likely composed of gp181.

Table 5.2: Comparative analysis of proteins involved in infection process.

KTN4 proteins	Predicted function	φKZ homologs	Amino acid sequence similarity	General characteristics
gp39	tail sheath protein	**gp29**	100%	*Tail apparatus*
gp40	tail tube protein	**gp30**	100%	
gp166	tail tip protein	**gp131**	99%	
gp167	structural protein, φKZ gp131 paralog	**gp132**	99%	
gp168	structural protein, φKZ gp131 paralog	**gp133**	99%	
gp169	structural protein, φKZ gp131 paralog	**gp134**	99%	
gp170	structural protein, φKZ gp131 paralog	**gp135**	99%	
gp181	tail sheath protein	**gp145**	99%	
gp182	tail fiber	**gp146**	98%	
gp180	endolysin	**gp144**	99%	*Lysis cassette*
gp221	peptidoglycan hydrolase	**gp181**	99%	
gp106	β/β′-like RNAP	**gp80** [a]	99%	*RNA polymerase*
gp185	β/β′-like RNAP	**gp149** [a]	99%	
gp219	β/β′-like RNAP	**gp178** [a]	99%	
gp220	β/β′-like RNAP	**gp180** [a]	99%	
gp74	nvRNAP	**gp55**	Identical, but shorter by 41 amino acids	
gp95	nvRNAP	**gp71**	99%	
gp96	nvRNAP	**gp73**	100%	
-	nvRNAP	**gp72**	-	
gp97	nvRNAP	**gp74**	100%	
gp158	nvRNAP	**gp123**	100%	
gp90	unknown function	**gp68**	100%	
gp04	dihydrofolate reductase	**gp04**	98%	*Nucleotide metabolism*
gp229	thymidylate kinase	**gp188**	100%	
gp258	deoxycytidine triphosphate deaminase	**gp214**	100%	
gp284	thymidylate synthase	**gp235**	99%	
gp367	ribonucleosidediphosphate reductase, beta chain	**gp305**	99%	
gp368	ribonucleoside-diphosphate reductase, alpha chain	**gp306**	100%	
gp52	chain A of monomeric subunit of Tubz	**gp39**	100%	

KTN4 proteins	Predicted function	φKZ homologs	Amino acid sequence similarity	General characteristics
gp98	Helicase	**gp75**	99%	*DNA replication and modification*
gp244	Helicase	**gp203**	100%	
gp114	prosite ATP-dependent DNA ligase	**gp87**	99%	
gp191	RNase H	**gp155**	100%	
gp76	intron-encoded endonuclease	**gp56**	99%	
-	intron-encoded endonuclease	**gp179**	-	
-	intron-encoded endonuclease	**gp296**	-	

β/β′-like RNAP - β/β′-like virion-associated non-canonical multi-subunit viral RNA polymerase
nvRNAP non-virion RNA polymerase

DNA replication and modification

The analysis revealed that gp52 is homologus to the chain A of monomeric subunit of Tubz protein of φKZ (φKZ gp39, 100% similarity, ESI-MS/MS hit). Aylett *et al.* structural studies suggest, that this protein preasumably is essential for the correct centering of replicated bacteriophage virions within the bacterial host (Aylett *et al.*, 2013).

Moreover, four β/β′-like virion-associated proteins in KNT4 phage (gp106, gp185, gp219, and gp220, ESI-MS/MS hit) were assigned as a non-canonical multi-subunit viral RNA polymerase (RNAP), 99% similar to φKZ gp80, gp149, gp178, gp180 (Ceyssens *et al.*, 2014) (Table 5.2). Furthermore, the analysis of φKZ RNAP performed by (Yakunina *et al.*, 2015) allowed to identify five homologous subunits in KTN4 genome, the products of early phage genes (Table 5.2). Four of these are cellular RNAP subunits homologs of the non-virion set (nvRNAP): 1) gp74, homologous to gp55 φKZ; 2) gp95 and gp96 (φKZ gp71 and gp73); φKZ gp72, a part of φKZ subunit has no homologue in KTN4 genome; 3) gp97 identical to gp74 φKZ and 4) gp158 identical to gp123 φKZ. The fifth subunit, gp90 identical to gp68 φKZ, is a protein of unknown function with no similarity to known RNAP subunits or any other known protein family. Gp74 and gp97 together correspond to msRNAP largest (bacterial β') subunits and gp95 and gp96 subunit corresponds to the C-terminal half of msRNAP second largest (β in bacteria) subunits. Gp158 is a highly diverged homolog of the N-terminal half of the second largest (β in bacteria) msRNAP subunits. This complex initiates transcription from the late promoters in rifampicin-resistant manner, tested *in vitro* in φKZ, which suggests that the virus relies on its own transcription machinery for the entire infection process. However, the late promoter 5'-TATG-3' conserved motif is necessary for transcription by nvRNAP *in vitro* (Yakunina *et al.*, 2015).

Furthermore, genes associated with nucleotide metabolism (KTN4 gp04, dihydrofolate reductase; gp229 thymidylate kinase; gp258, deoxycytidine triphosphate

deaminase; gp284, thymidylate synthase; gp368, ribonucleoside-diphosphate reductase, alpha chain, and gp367, ribonucleosidediphosphate reductase, beta chain) were identified. All are identical or highly similar to φKZ genes: gp04 (98%), gp188 (100%), gp214 (100%), gp235 (99%), gp306 (100%) and gp305 (99%), respectively (Mesyanzhinov *et al.*, 2002).

Also proteins playing role in DNA replication (KTN4 gp98 and gp244, a potential helicase; gp114, a prosite ATP-dependent DNA ligase AMP-binding site pattern, gp191, a putative RNase H have homologs in φKZ genome [gp75 (99%), gp203 (100%), gp87 (99%) and gp155 (100%), respectively] (Mesyanzhinov *et al.*, 2002).

However, when intron-encoded endonucleases (Mesyanzhinov *et al.*, 2002) were examined significant variations between both phages were observed. The φKZ gp179 and gp296 are missing homologs in KTN4 genome and only gp56 is 99% similar to KTN4 gp76.

5.3.4. Comparative genome analysis and protein-sharing network

The protein sharing network for jumbo phage KTN4 is presented in Fig. 5.14. A resulting network comprises 495 phages (nodes) belonging to *Myoviridae*, *Siphoviridae*, *Podoviridae*, or uncharacterized and other phages and 6,948 relationships (edges) between them (Fig. 5.14 A). In this graph, phage KTN4 was placed in a single component with nine well-known giant phages including φKZ, φPA3, 201φ2-1, EL, OBP, ΦJM-2012, SPN3US, CR5 and PhiEaH2. When the network topology was computed with two classical measures such as the clustering coefficient (CL) and betweenness centrality (BC) (Brohée *et al.*, 2008), this component shows the highest CL = 1 (absolute cohesiveness) and the lowest BC = 0 (none of node acting as a bridge among other pairs of nodes). This network structure reflects the distinct core gene-sets shared between their genomes (Jang *et al.*, 2013), which form a tight-knit clique of full interconnectivity. Next, the connectivity pattern of this component has been investigated not only on the basis of protein sequence identities, but also according to phage-phage similarity score after normalizing the number of shared genes between genomes (Leplae *et al.*, 2010; Lima-Mendez *et al.*, 2008). As a result, over the threshold of 60% identity, phages KTN4, φKZ, phiPA3, and 201φ2-1 kept forming an in-group relationship with φKZ being the closest relative to KTN4 (Fig. 5.14 B). In addition, KTN4 was more closely related to φKZ, phiPA3, and 201φ2-1 with the phage-phage similarity score ranging from 157.2 to 999.9, indicating more shared homologous genes between them than others. The other

Pseudomonas phages including EL and OBP were connected for identity values less than 35% and similarity score, 79.9. The phages PhiEaH2, SPN3US, CR5, and ΦJM-2012, having different host ranges (i.e., *Erwinia*, *Salmonella*, *Cronobacter*, and *Vibrio sp.*, respectively), were weakly connected to other φKZ-related members in terms of both shared gene contents and sequence identity. The genetic relationships of KTN4 have been investigated by constructing a mathematical model of gene (protein)-sharing network, extending to possible close relatives. In our phage population network, KTN4 is constricted to a single isolated component comprising the four φKZ-related phages and other potential relatives. Subsequent network decomposition strongly indicates that KTN4 belongs to the *Phikzvirus* (Cornelissen *et al.*, 2012), with the large proportion of φKZ specific core gene-sets in common to the φKZ, φPA3, and 201φ2-1. More specifically, the connectivity patterns suggest that φKZ appears to be the closest relative to KTN4 as their protein families with more than 90% sequence identity can be considered more recently shared than those of other phage members in this group (Halary *et al.*, 2010). In addition, our population network can reveal other informative connections. The phages phiEaH2, SPN3US, and CR5 were found to interconnect solely with the φKZ-related phages, indicating that they are probably diverged member of the φKZ-related group as observed in ΦJM-2012 (Jang *et al.*, 2013). These results suggest that the *Phikzvirus* phages' diversity has not been fully delimited and that there are additional more distant relatives yet to be discovered.

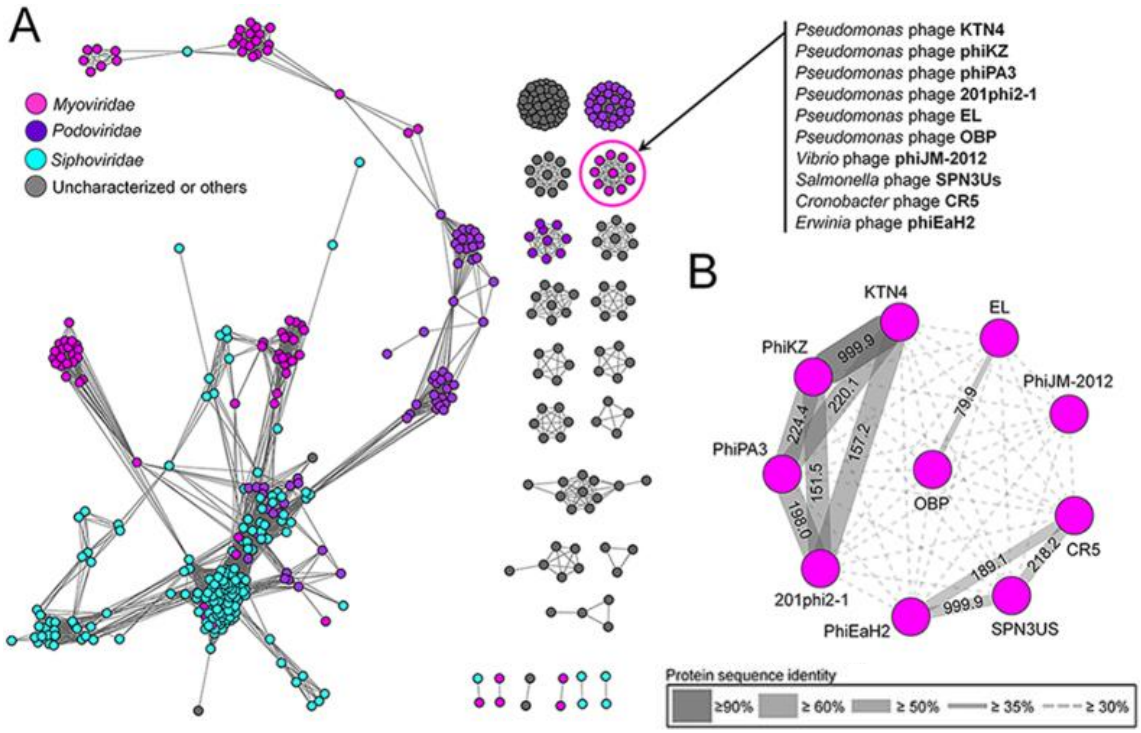


Figure 5.14: Protein-sharing network for KTN4. (A) A network representation was produced using the edge-weighted spring embedded layout of Cytoscape version 3.1.1. Nodes indicate phage genomes and edges between two nodes indicate their statistically weighted pairwise similarities with phage-phage similarity scores of ≥ 1. There are 495 nodes and 6,948 edges in this network. (B) An enlarged view of the circle in Panel B. Values are the similarity scores estimated with the hypergeometric equation shown in Materials and methods. Edge thickness is proportional to protein sequence identity, which is represented in the legend box.

The general features of KTN4 phage are presented in table 5.3.

Table 5.3: Major features of KTN4 phage.

Phage	**KTN4**
Water source	Irrigation field
City/Country/ Date	Wrocław/Poland/ 2011
Host	PAO1
Genome size (bp)	279,593
GC (%)	36.9
ORFs	368
Unique ORFs	19
Virion dimensions head/tail [nm]	130/168
Latent period [min]	40
Burst size [pfu/cell]	8±2
Heat stability for 1h [ºC]	40-70
pH stability for 1h at RT	6-12
Chloroform sensitivity at RT/4 ºC [h]	24
Pyocyanin production inhibition in biofilm structure [a]	+
Pyoverdin production inhibition in biofilm structure [a]	+
Increase of diffusion though biofilm structure [a]	+

[a]Experiments performed by Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University in Kielce, Kielce, Poland; Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Wrocław, Poland (Danis-Wlodarczyk *et al.*, 2015).

5.4. Discussion

In this chapter we report the discovery and characterization of a giant virus KTN4, a φKZ isolate, that belongs to the *Myoviridae* family and infects *P. aeruginosa.*

Microbiological characteristics

KTN4 has an enlarged capsid and a long contractile tail (130 nm and 168 nm, respectively). It uses the IV type fimbriae as receptor on bacterial surface. KTN4 was stable in a broad range of temperature (40 - 70°C), pH (6 - 12) and was resistant to chloroform treatment. This lytic virus has a broad spectrum of activity with prevalence to clinical isolates, especially from CF patients (36.8% and 32.6% in two independent *P. aeruginosa* panels). However, a highly similar phage φKZ exhibited broader host range, lysing 37% and 46%. The detailed comparison of KTN4 and φKZ genes have been done, especially for tail fibers and its C-terminal elements as responsible for host receptor recognition. We have looked also for differences in auxiliary genes conditioning for instance DNA modification which may result in phage DNA resistance to host restriction digestion mechanism. There was no differences in tail fibers gene sequences. However, the φKZ genome possess several unique genes (no. 17) in the area of auxiliary genes such as hypothetical proteins surrounding terminase large subunit, non-virion DNA- dependent RNA polymerase subunit, DnaB helicase, similar to putative endonuclease (*Pseudomonas* phage φPA3), cyclic GMP-AMP synthase; DNA sensor, transferase, putative dimetal phosphatase; dinuclear metal center phosphatase, metalloprotein, metallophosphoesterase, endonuclease S subunits; coiled-coil, type I methyltransferase, DNA binding and homing endonuclease. The comperative analysis suggest, that there are no differences in host receptor recognition, destabilization of bacterial peptidoglycan as well as DNA metabolism and RNSP. Potentially the host range discrepancy could be related with variations between replication and DNA modification genes.

One-step experiment revealed a latent period of 40 min and a burst size of 8 ± 2 phage particles per infected cell. *P. aeruginosa* PAO1 cell infection cause by KTN4 was examined in the real time with the use of time-lapse microscopy. In the early stage of infection the small circular mass was observed, which migrates toward the cell midpoint and increases in size. It is likely the site of genome encapsidation as it was previously demonstrated by Kraemer *et al.* (2012). After arrival at midcell, the mass oscillates slightly about the central axis. Erb *et al.* (2014) hypothesized that this

oscillatory movement is reminiscent of eukaryotic chromosomes lined up at midcell by the mitotic spindle. In contrast to eukaryotes, where DNA replication is temporally separated from segregation, the phage nucleoid continued to increase in mass as it moved toward midcell, suggesting that DNA replication and positioning occur simultaneously. This observation demonstrates that bacteriophage have possibility to harness the properties of a tubulin-like cytoskeleton for efficient propagation. In further step of infection, bacteria cell continues to increase in size to finally achieve a significantly enlarged lemon-like shape. Erb *et al.* (2013 and 2014) observed very similar observation during *P. chlororaphis* cells 201ϕ2-1 phage infection. In their study a long fibers of PhuZ, a type of tubulin, form and arrange into a structure that resembles a simple form of the spindle apparatus. These structures are only formed if a virus has invaded a cell. During the early stages of infection, viral DNA localizes to the cell poles, but rearranges at around the same time that PhuZ fibers form. At this point, the viral DNA moves towards the middle of the bacterial cell with fibers on either side of the DNA, suggesting that the fibers play a role in moving the DNA. Kraemer *et al.* (2012) also hypothetized, that PhuZ tubulins play a role in spatially organizing DNA during lytic growth and thereby contribute to efficient phage production. This bacteriophage system may be an evolutionary ancestor of the spindle apparatus. In conclusion, bacteriophage can exploit the advantages of a well-defined cytoskeletal organization to faithfully propagate themselves. The PhuZ spindle appears to play a key role in organizing viral reproduction by positioning the phage nucleoid at midcell. Since polar localization is phage infection-dependent, Erb *et al.* (2014) hypothesized that a phage protein is necessary for the organization of the spindle. Whether the putative phage polar protein also acts to nucleate *de novo* filament formation or facilitate bundling is unclear.

Biofilm eradication ability

Since *Pseudomonas* common infections are usually associated with biofilm formation, the ability of KTN4 to disrupt the biofilm has been examined. Initial evaluation was performed on 24 h and 48 h old *P. aeruginosa* PAO1 biofilms with the use of confocal scanning microscopy. Untreated biofilm presented a dense structure, which thickness increased over the time. However, after phage treatment biofilm structure was less dense with clear areas of biofilm removal. Further evaluation with more qualitative and quantitative techniques was performed by Drulis-Kawa group (Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University

of Wrocław, Wrocław, Poland) and collaborators (Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University in Kielce, Kielce, Poland and Zeta Instruments Co., San Jose, USA). Results are published in Danis-Wlodarczyk *et al.*, 2016.

The analysis of pyocyanin and pyoverdin/pyochelin secretion to the medium was carried out using spectrophotometry and fluorometry, allowing detection of highly diffusible pigmented signaling molecules levels of quorum sensing and a siderophor. The analysis of pyocyanin concentration in growth medium showed that active phages, colistin and active phage/colistin treatment significantly decreased the level of this compound for the tested biofilms (24, 48 or 72 h) compared to UV-inactivated phages. It turned out that the infective form of our giant phage significantly reduced the concentration of pyocyanin, whether combined with colistin or alone. The level of pyoverdin determined by the fluorescence was significantly lowered after colistin application on 24 h biofilm and in combination with both phage preparations on PAO1 biofilm formed for 72 h. The reduction of siderophore concentration was also observed when 72 h biofilms were exposed to active KTN4. The combined treatment consisting of phage and antibiotic did not showed a synergistic effect in either biomass or dyes determination assays. Statistically significant ($p < 0.005$) inhibition observed for pyocyanin and pyoverdin/pyochelin production was elicited by phages (lysis of cells in biofilm). The quantities of the most important pigmented signaling molecules secreted by the pathogen decreased significantly, which proves the potency of these phages to be applied in *Pseudomonas* biofilm treatment.

The phage application effect on biofilm disruption was analyzed also by laser interferometry, a biophysical technique which measures quantitative changes in biofilm matrix permeability for low molecular compounds. The diffusion of TSB medium though biofilm was evaluated, which indirectly indicated its structure degradation. The hydrophilic Nephrophane membrane overgrown by 72 h biofilm was chosen for the interferometry analysis. The PAO1 biofilm was treated with active and UV-inactivated KNT4 phage for 4 h. The diffusion rate of medium transported though the biofilm-covered membrane (0.86×10^{-3} mg/h) was significantly higher than for intact biofilm ($p < 0.001$) after active and inactivated phage treatment, reaching 1.65×10^{-3} mg/h and 1.39×10^{-3} mg/h, respectively. The increase of diffusion rate through the overgrown membrane after phage application indicated the degradation of the biofilm structure associated with the disintegration of matrix elements.

Airway Surface Layer Infection model

A strong antibacterial activity (4 - 7 logs decrease in extracellular bacterial load) was observed in an infected Airway Surface Layer Infection model using NuLi-1 and CuFi-1 cell lines. This assay also showed that KTN4 successfully prevented wild-type *Pseudomonas* internalization to epithelial cells. These properties indicate giant KTN4 phage as an excellent candidate for phage therapy purposes and is a source for antimicrobials and biotechnologically relevant proteins. To our knowledge this is the first study of phage application using this lung epithelia infection assay. The gentamicin exclusion assay on ASL *in vitro* model is flexible, generates reproducible data with well-controlled and standardized conditions, mimicing the normal and CF lung environments. Moreover, it provides a basis for understanding the host-pathogen interactions and is, as such, an important step towards experimental *in vivo* studies.

Genome properties

Phage KTN4 has a genome size of 279,593 bp with 368 predicted ORFs and 6 tRNAs. It shows a genome-wide nucleotide sequence similarity to: φKZ 99%, PA7 99%, φPA3 84%, 201φ2-1 78% (BLASTN). As such, it can be defined as an isolate of the *Pseudomonas* phage φKZ species. The average GC content of KTN4 (36.9% GC) is significantly lower than that of the host *P. aeruginosa* (66% GC). The tRNAs are not conserved among the other φKZ-related phages, which also have a lower GC content [φKZ (36.8% GC), EL (49.3%), and 201φ2-1 (45.3%)] than the respective *Pseudomonas* species they infect. The absence of notable homology to other bacteriophages on the DNA level indicates clearly that KTN4 represents an evolutionarily distinct group in the *Myoviridae* family. In total, only 21.5% of all predicted proteins have defined function and ESI-MS/MS analysis allowed identification of 111 virion-associated proteins. According to the orientation of transcription, ORFs are organized into operons and most are on a positive strand. Moreover, 47 promoters and 107 putative factor-independent terminators were predicted.

Lysis cassette

The KTN4 endolysin (gp180) is highly identical (99%) to the endolysin of *P. aeruginosa* φKZ gp144. This enzyme is well studied both at the molecular, biochemical (Briers *et al.*, 2007, 2009; Paradis-Bleau *et al.*, 2007; Cloutier *et al.*, 2010)

and structural level (Fokine *et al.*, 2008). φKZ gp144 is one of the first reported modular endolysins active against Gram-negative peptidoglycan (chemotype A1γ) (Briers *et al.*, 2007). It comprises a N-terminal peptidoglycan binding domain that has high binding affinity for its substrate (295 nM) (Briers *et al.*, 2009) and a C-terminal transglycosylases domain that cleaves the glycosidic linkage between N-acetylmuramic acid and N-acetylglucosamine with the concomitant formation of 1,6-anhydromuramic acid (Paradis-Bleau *et al.*, 2007). The crystal structure has been determined for free gp144 (2.5 Å) and complexed with chitotetraose (2.6 Å) (Fokine *et al.*, 2008). The consensus motifs for peptidoglycan binding (underlined) and the catalytic residue (boxed) are fully conserved in KTN4 gp180. Whereas Briers *et al.* (2007) could purify high amounts of recombinant gp144 from the cytosolic fraction, Paradis-Bleau *et al.* (2007) found that φKZ gp144 expression results in host cell lysis and that the endolysin is also secreted independent of a holin, the Sec (or TAT) secretion machinery. Bacterial membrane permeabilizing activity has been demonstrated for φKZ gp144 using fluorescence spectroscopy, putatively initiated by its weakly positively charged N-terminus (Cloutier *et al.*, 2010). Extensive bio-informatical analysis did not predict any typical holin in φKZ. In line with these findings, no clear holin sequence could be predicted in KTN4 as well. It remains to be elucidated how *Phikzvirus* phages achieve lysis in *in vivo* conditions, and if the lysis system differs from previously described systems such as the canonical lambda lysis system (Young *et al.*, 2000), endolysins with a signal peptide (São-José *et al.*, 2000) or a signal-arrest-release sequence (Xu *et al.*, 2004).

Moreover, the ESI-MS/MS analysis identified the presence of tail associated enzyme gp221 (structural peptidoglycan hydrolase), that correspond to gp181 of φKZ phage (99% similarity). This enzyme cleaves the host cell wall during the first stage of the life cycle (Briers *et al.*, 2008). In the centre of the φKZ baseplate, there is a density that resembles the needlelike ‘‘cell-puncturing’’ device of T4, which is most likely composed of gp181.

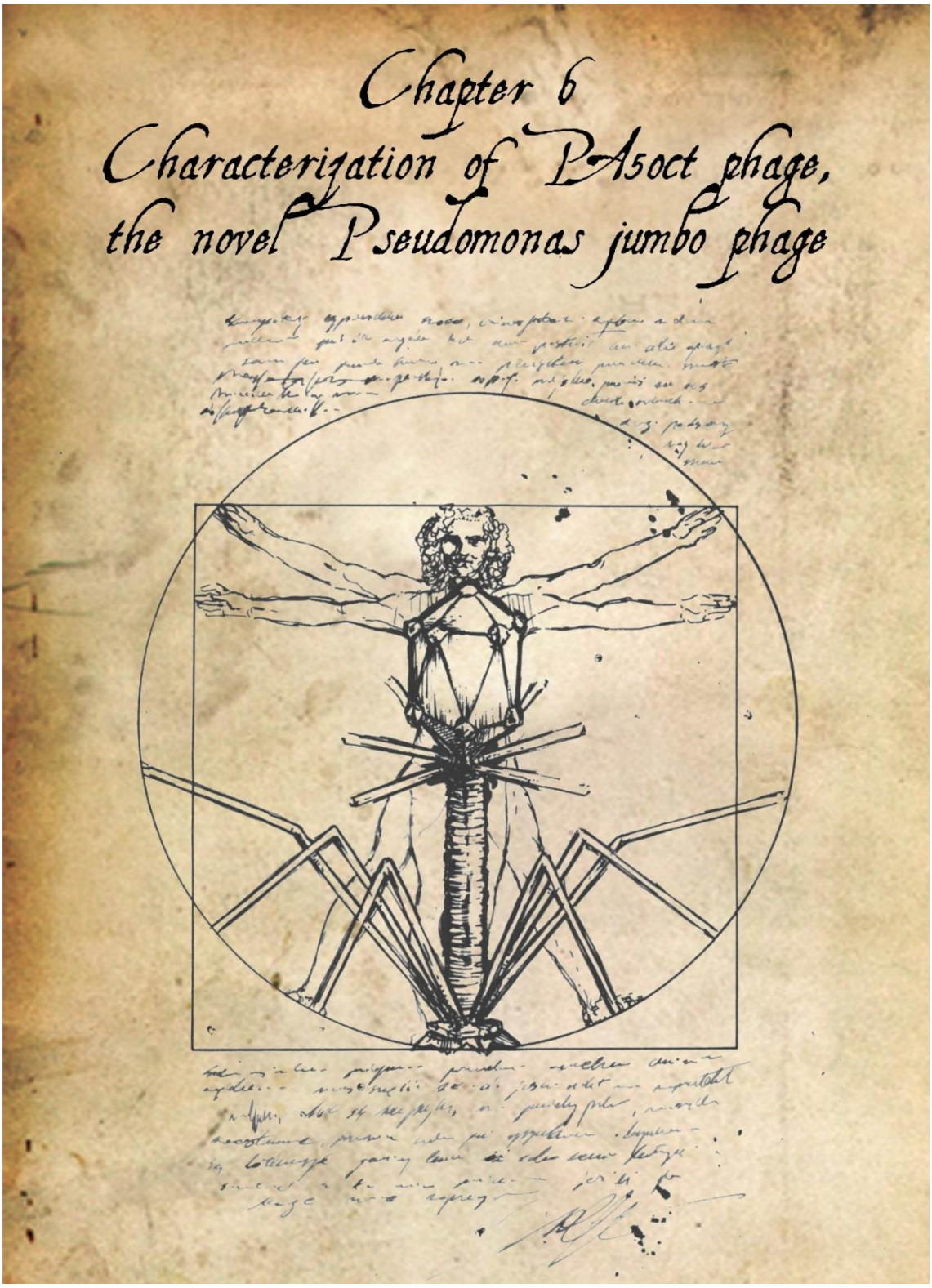
Protein sharing network

The genetic relationships of KTN4 have been investigated by constructing a mathematical model of gene (protein)-sharing network, extending to possible close relatives. In our phage population network, KTN4 is constricted to a single isolated component comprising the five *Pseudomonas* φKZ related phages and other potential relatives. Subsequent network decomposition strongly indicates that KTN4 belongs

to the *Phikzvirus* clade (Cornelissen *et al.*, 2012), with the large proportion of φKZ-specific core gene-sets in common to the φKZ, φPA3, and 201φ2-1. More specifically, the connectivity patterns suggest that φKZ appears to be the closest relative to KTN4 as their protein families with more than 90% sequence identity can be considered more recently shared than those of other phage members in this group (Halary *et al.*, 2010). In addition, our population network can reveal other informative connections. The phages phiEaH2, SPN3US, and CR5 were found to interconnect solely with the φKZ-related phages, indicating that they are probably diverged member of the φKZ-related group as observed in ΦJM-2012 (Jang *et al*., 2013). These results suggest that the diversity of *Phikzvirus* phages has not been fully delimited and that there are additional more distant relatives yet to be discovered.

Considering all these characteristics, KTN4 phage is a suitable and promising candidate for *in vivo* trials, for applications in treatment and prophylaxis in lung infections.

Chapter 6
Characterization of PA5oct phage, the novel Pseudomonas jumbo phage



Contribution

- Transmission electron microscopy (TEM) phage visualisation and morphology analysis was performed by Prof. Hans-W. Ackermann (Department of Microbiology-Infectiology and Immunology, Medical School, Laval University, Quebec, Canada) and technician Sylwia Nowak (University of Wrocław, Poland).
- The time-lapse microscopy experiment was conducted in the Centre for Food and Microbial Technology (Prof. Abram Aertsen, KU Leuven, Belgium) with a help of Dr. William Cenens.
- ASL and biofilm experiments were performed during research visit at the Department of Molecular Medicine, Royal College of Surgeons in Ireland, Education and Research Centre, Beaumont Hospital, Dublin, Ireland (Prof. Brian J. Harvey) with a help of Dr. Gerard Higgins and Dr. Jean Tyrrell.
- The protein-sharing network was constructed and analyzed by Dr. Ho Bin Jang (Laboratory of Gene Technology, KU Leuven, Belgium) and Katarzyna Danis-Włodarczyk.
- RNA seq experiments were performed with the support of Bob Blasdel (Laboratory of Gene Technology, KU Leuven, Belgium).
- Promoters identification performed with the support of Dr. Dieter Vandenheuvel (Laboratory of Gene Technology, KU Leuven, Belgium).
- ESI-MS/MS was conducted in Biomedical Research Institute and Transnationale Universiteit Limburg, School of Life Sciences, Hasselt University, Diepenbeek, Belgium (Prof. Jean-Paul Noben).
- All remaining experiments and analysis were performed by Katarzyna Danis-Włodarczyk.

6.1. Introduction

Viruses have the genome represented by either RNA or DNA, that can be either single-stranded or double-stranded, either circular or linear, and consists of either a single or multiple molecules (Agol, 1974; Baltimore, 1971). The genomes are typically smaller compared to genomes of cellular life forms. However, over the past few years the discovery of several groups of giant/jumbo viruses has dramatically expanded the viral genome size range that now spans three orders of magnitude, from about 2 kb to over 2 Mb. Often they are larger than the genomes of numerous bacteria and archaea, obliterating the gulf between cells and viruses in terms of genome size and complexity (Claverie *et al.*, 2009; Koonin *et al.*, 2015).

To our knowledge, 34 *Myoviridae* and 9 *Siphoviridae* jumbo phages have been described and sequenced to date (Table 6.1). These include seven *Pseudomonas* phages, EL (211 kbp, Krylov *et al.*, 2003), φKZ (280 kbp, Mesyanzhinov *et al.*, 2002), phiPA3 (309 kbp, Monson *et al.*, 2011), 201phi2-1 (316 kbp, Thomas *et al.*, 2008), PaBG (258 kbp, Sykilinda *et al.*, 2014), Lu11 (280 kbp, Adriaenssens *et al.*, 2012) and OBP (284 kbp, Cornelissen *et al.*, 2012).

Table 6.1: Representatives of sequenced jumbo bacteriophages (NCBI).

Phage	Accession no.	Length	Protein no.	Host
Myoviridae				
Phikzvirus				
φKZ	NC_004629	280334 nt	306	*Pseudomonas*
Elvirus				
EL	NC_007623	211215 nt	201	*Pseudomonas*
Rsl2virus				
RSF1	NC_028899	222888 nt	230	*Ralstonia*
RSL2	NC_028950	223932 nt	224	*Ralstonia*
Rslunavirus				
RSL1	NC_010811	231255 nt	343	*Ralstonia*
Spn3virus				
SPN3US	NC_027402	240413 nt	264	*Salmonella*
Agrican357virus				
Ea35-70	NC_023557	271084 nt	318	*Erwinia*
Unclassified T4-like phages				
PX29	NC_023688	222006 nt	330	*Aeromonas*
phiAS5	NC_014636	225268 nt	343	*Aeromonas*
CC2	NC_019538	231743 nt	427	*Aeromonas*
Aeh1	NC_005260	233234 nt	352	*Aeromonas*
65	NC_015251	235229 nt	437	*Aeromonas*
phiN3	NC_028945	206713 nt	402	*Sinorhizobium*
ACG-2014f	NC_026927	228143 nt	292	*Synechococcus*
KVP40	NC_005083	244834 nt	381	*Vibrio*
VH7D	NC_023568	246964 nt	327	*Vibrio*
nt-1	NC_021529	247511 nt	405	*Vibrio*
P-SSM2	NC_006883	252401 nt	334	*Prochlorococcus*
unassigned Myoviridae				
G	NC_023719	497513 nt	675	*Bacillus*
unclassified Myoviridae				
0305φ8-36	NC_009760	218948 nt	246	*Bacillus*
φPA3	NC_028999	309208 nt	375	*Pseudomonas*
201φ2-1	NC_010821	316674 nt	461	*Pseudomonas*
PAU	NC_019521	219372 nt	295	*Sphingomonas*
CR5	NC_021531	223989 nt	231	*Cronobacter*
S-SSM7	NC_015287	232878 nt	319	*Synechococcus*
PaBG	NC_022096	258139 nt	308	*Pseudomonas*
phiR1-37	NC_016163	262391 nt	367	*Yersinia*
Lu11	NC_017972	280538 nt	391	*Pseudomonas*
OBP	NC_016571	284757 nt	309	*Pseudomonas*
vB_KleM-RaK2	NC_019526	345809 nt	534	*Enterobacteria*
K64-1	NC_027399	346602 nt	64	*Klebsiella*
PBECO 4	NC_027364	348113 nt	551	*Escherichia*
121Q	NC_025447	348532 nt	611	*Escherichia*
vB_CsaM_GAP32	NC_019401	358663 nt	545	*Cronobacter*
Siphoviridae				
Phicbkvirus				
phiCbK	NC_019405	215710 nt	338	*Caulobacter*
Magneto	NC_019407	218929 nt	347	*Caulobacter*
Swift	NC_019411	219216 nt	343	*Caulobacter*
karma	NC_019410	221828 nt	353	*Caulobacter*
Rogue	NC_019408	223720 nt	350	*Caulobacter*
unclassified Siphoviridae				
phiEaH1	NC_023610	218339 nt	241	*Erwinia*
phiEaH2	NC_019929	243050 nt	262	*Erwinia*
CcrColossus	NC_019406	279967 nt	448	*Caulobacter*
S-SKS1	NC_020851	208007 nt	281	*Synechococcus*

However, there is a high possibility, that jumbo phages are more abundant in the environment and have important impact on bacteria. Difficulties of their isolation and identification are due to the fact that some of them form very small plaques (< 0.5 mm on 0.7% soft agar), which are not visible or visible only in very low concentrations of soft agar (0.15 - 0.2%). This leads to their omission during classical phage propagation procedures and a major pool of them remains undetected. A good example is *Bacillus megaterium* phage G, the biggest isolated bacterial virus with genome size of 497,513 bp, discovered only by accident through electron microscopy of another bacteriophage preparation (Serwer *et al*., 2007).

The general knowledge of jumbo phages is limited, therefore it is very important to study their genome organization, mechanisms of evolution and biology (Hendrix, 2009). In this chapter we present a new jumbo myovirus, PA5oct lytic against *P. aeruginosa.* We focus on the biology, genome structure, proteome, evolutionary relationships and possibilities of application in phage therapy to combat infections caused by *P. aeruginosa*.

6.2. Microbiological characteristics

6.2.1. Isolation and morphology

PA5oct is a novel jumbo *Pseudomonas* phage isolated from a sewage sample collected from irrigated fields near Wrocław, Poland. After purification phage titres were 10^8 - 10^{10} pfu/ml and caused < 0.5 mm wide clear plaques on 0.7% soft agar. PA5oct was classified to order *Caudovirales,* family *Myoviridae* (Drulis-Kawa *et al.,* 2013; Lavigne *et al.,* 2009) and formally named vB_PaeM_PA5oct (PA5oct). TEM images analysis revealed, that heads measured 131 nm between opposite apices (60 measurements) and were icosahedral as shown by the simultaneous observation of capsids with hexagonal and pentagonal outlines. Tails measured 136 x 19 nm in the extended state and consisted of a neck of 10 nm, a sheath with about 25 transverse striations, and an indistinct baseplate of 33 x 7 - 10 nm with short fibers of 30 x 2 nm. No collar was observed, but many phages carried amorphous material at the neck. Contracted tail sheaths measured 50 x 30 nm and were separated from the baseplate by a space of about 2 nm (Fig. 6.1).

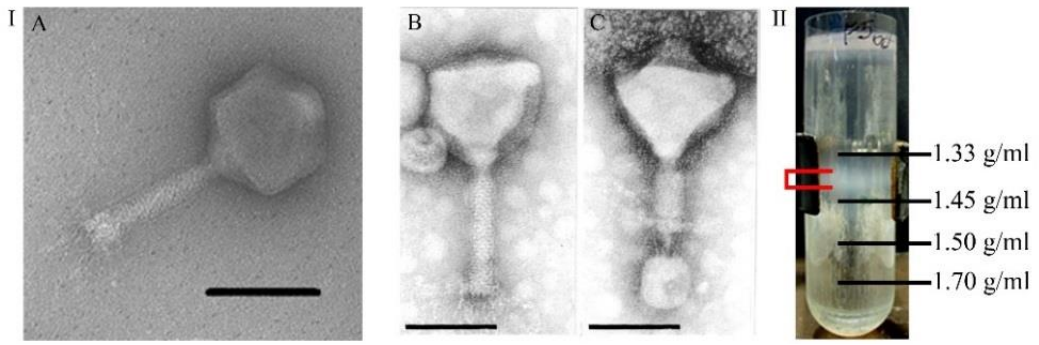


Figure 6.1: Particle analysis of PA5oct phage. (I) an electron microscopic images with the scale bar representing 100 nm, uranyl acetate, final magnification 297,000 ×; (A) and (B) phage with pentagonal head and an extended tail, (C) phage with a contracted tail; (II) a CsCl gradient purification (four different CsCl densities are marked: 1.33 g/ml, 1.45 g/ml, 1.50 g/ml, 1.70 g/ml; phage band pronounced red).

6.2.2. Phage infection

One-step growth experiments indicated a latent period of 40 min and a burst size of about 30 - 40 phage particles per infected bacterial cell (Fig. 6.2).

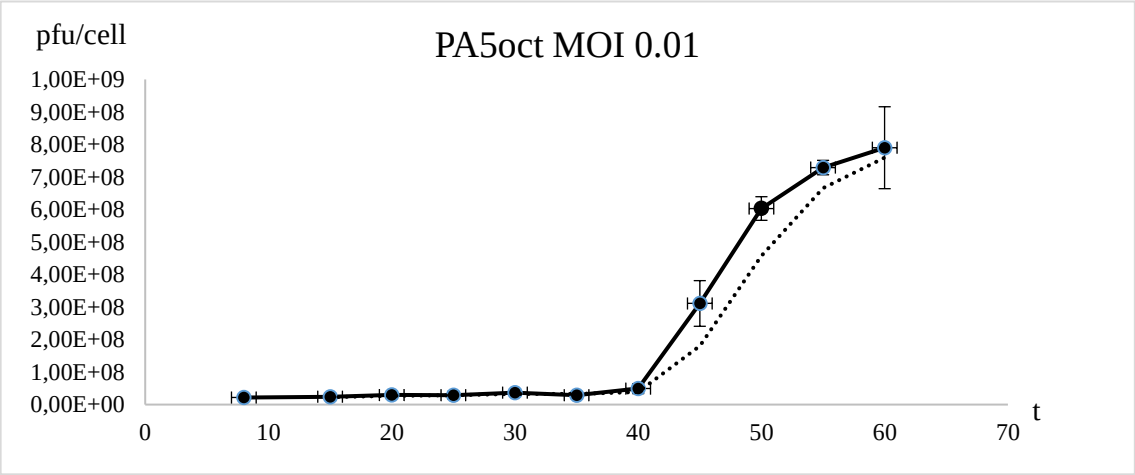


Figure 6.2: One-step growth curve of phage PA5oct. An equal volume of bacterial culture (at OD_{600} of 0.4, 10^8 cfu/ml) was mixed with phage suspension (10^6 pfu/ml) to obtain a MOI of 0.01. A trend line is marked with dots. All data are collected from three independent experiments, error bars indicate the standard deviation.

6.2.3. Time-lapse microscopy

The *P. aeruginosa* cells infection caused by phage PA5oct in the real time was visualized with a time-lapse microscopy (Eclipse Ti Time-Lapse Microscope). *P. aeruginosa* PAO1 Krylov cells were grown to mid-exponential phase (OD_{600} of 0.6) and infected on a 0.2% PBS-agar pad by CsCl purified PA5oct phage (9.79 x 10^9 pfu/ml) with MOI 10. In the course of infection cells only marginally increases their size and after 40 - 45 min from the start of infection bacterial cells burst and phage progeny was released (Fig. 6.3). The latent period of PA5oct achieved in this experiment cannot be compared to one step growth experiment, due to different conditions.

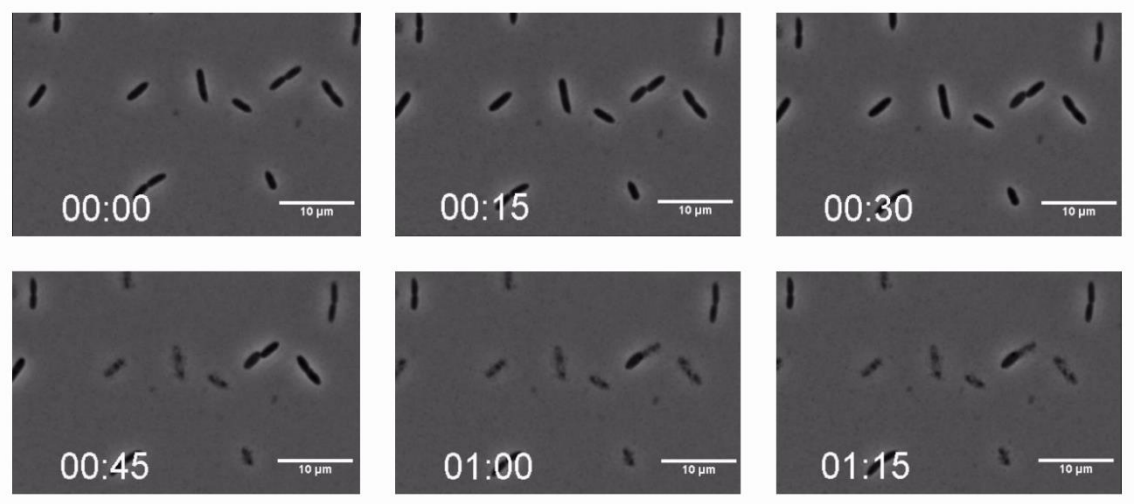


Figure 6.3: Time-lapse microscopy of the *P. aeruginosa* PAO1 Krylov infection caused by phage PA5oct on a 0.2% PBS-agar pad. A snapshots were taken from the start of phage infection every 15 min, MOI 10, using a Nikon Eclipse Ti Time-Lapse Microscope. Scale bare represents 10 µm.

6.2.4. Stability tests

Phage PA5oct (9.79 x 10^8 pfu/ml) was stable at temperatures from 4 - 60°C, where no reduction of pfu/ml was observed over a period of 60 min, while 5 min and over than 15 min incubation at 70°C decreased the titre by 4 and 6 logs, respectively (Fig. 6.4).

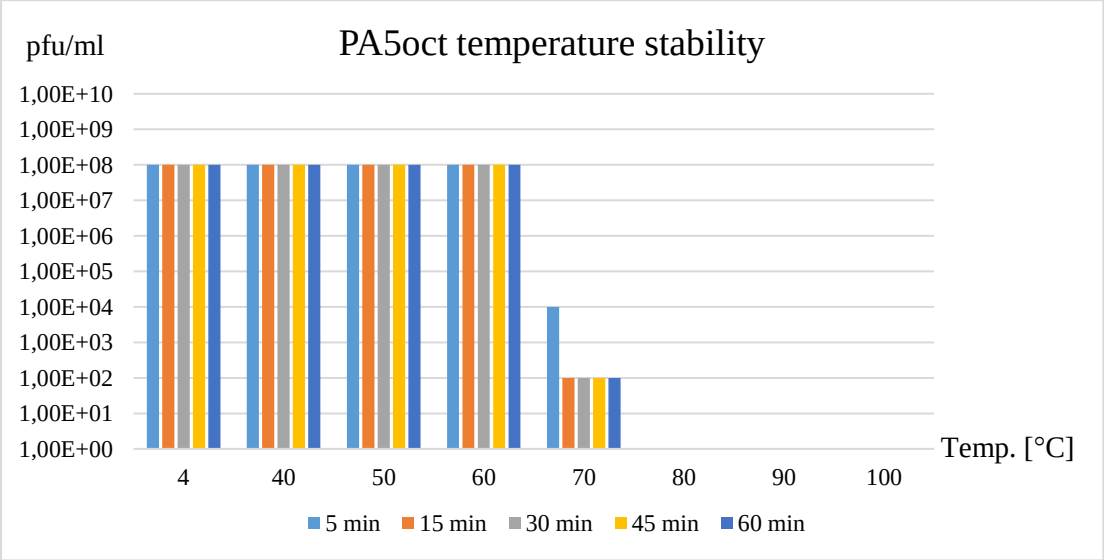


Figure 6.4: The temperature stability of phage PA5oct. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

Phage particles were also stable within a pH range of 5 - 11 after 1 h incubation at room temperature. When exposed to pH 4.0 and 3.0 for 1 h, phage titers dropped by 1 log, pH 2.0 and pH 12.0 by 3 and 4 logs, respectively (Fig. 6.5).

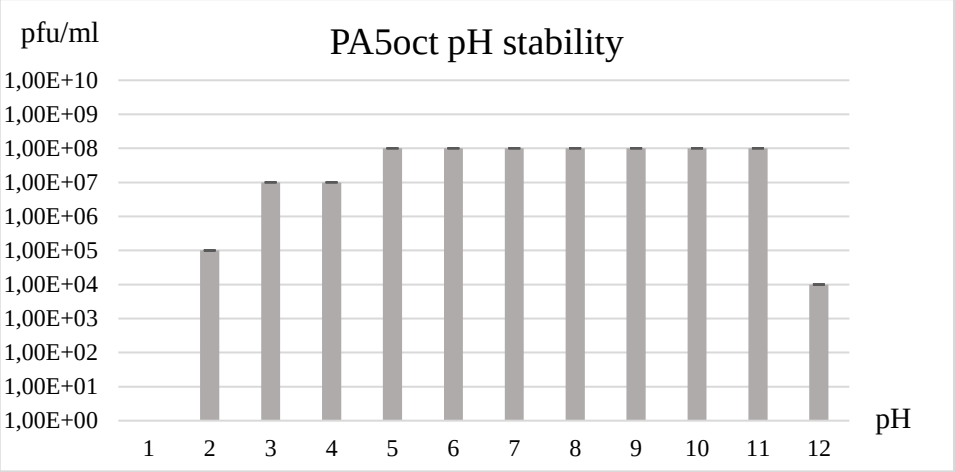


Figure 6.5: The pH stability of PA5oct phage. Phage titers (pfu/ml) were validated at the different pH values after 1h exposure at room temperature. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

No reduction in infectivity was observed after treatment with chloroform (an equal volume of phage solution and chloroform) for 1 h and 24 h at room temperature or at 4°C (Fig. 6.6).

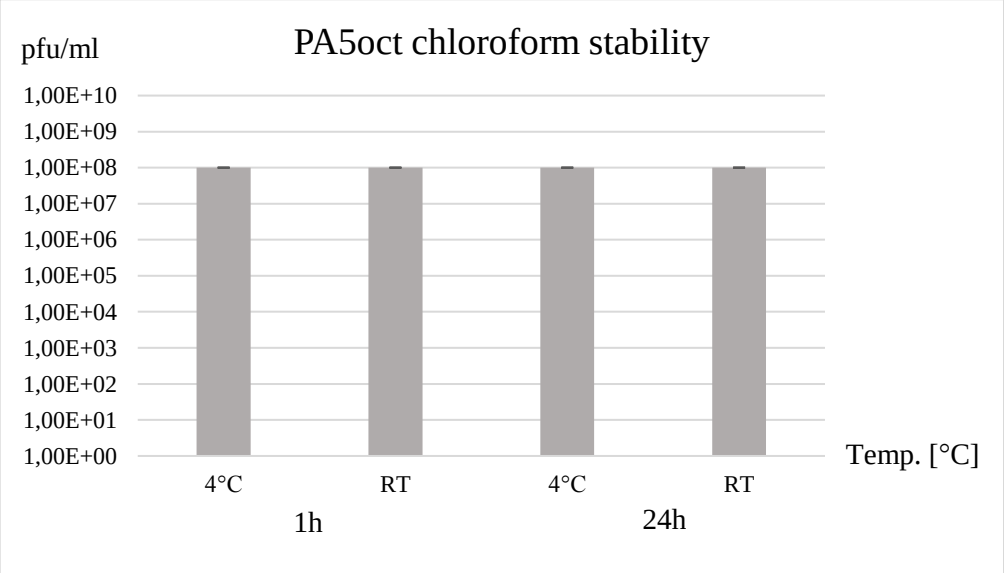


Figure 6.6: The chloroform stability of phage PA5oct. Phage titers (pfu/ml) were validated after chloroform treatment for 1 h and 24 h incubation at room temperature (RT) and at 4 °C. Each bar represents the mean of three independent experiments, error bars indicate the standard deviation. Results obtained with spot test.

6.2.5. Determination of phage receptor

Experiment with *P. aeruginosa* PAO1 mutants revealed that PA5oct possibly requires three receptors on bacterial cell surface: the type IV pili, LPS and flagella (Table 6.2). However, LPS mutations can have influence on overall makeup of the cell wall and how well flagella and pili are exposed on the bacterial surface. Good example

are Abeyrathne *et al.* (2005) studies of PAO1 *waaL* mutants. They proved, that this particular mutation drastically affected the swimming and twitching motilities of *P. aeruginosa*. The waaL encodes a functional O-antigen ligase, that has substantial influence on cell wall integrity and bacteria motility. For this reason, bacterial surface receptors of PA5oct should be further studied, by e.g. atomic force microscopy.

Table 6.2: Phage receptor identification on *P. aeruginosa* PAO1 mutants.

Bacterial strain	Phenotype	Origin	PA5oct
PAO1 (ATCC 15692)	Wild type	American Type Culture Collection	+
PAO1 Pirnay	Wild type with inactive type IV pili	Military Hospital Nederoverheembeek, Brussels, Belgium	-
PAO1 Krylov	Wild type		+
PAO1 ΔrmD (A-, B+)	Deficiency in D-rhamnose biosynthesis; lack of A-band LPS	Department of Molecular and Cellular Biology, University of Guelph, Canada	-
PAO1 ΔrmLC (A-, B-, core-)	Deficiency in L-rhamnose biosynthesis; truncate core region, lack of A-band and B-band LPS		-
PAO1 ΔwaaL (A-, B-)	Lack of WaaL ligating O-polymer to core-lipid A; LPS is devoid of A-band and B-band, semi rough (SR-LPS, or core-plus-one O-antigen)		-
PAO1 ΔwbpL (A-, B-)	Lack of glucosyltransferase WbpL essential for initiation of both A-band and B-band synthesis		-
PAO1 ΔfliC ΔalgC ΔpilA	Lack of flagella; lack of AlgC required for A-band, core oligosaccharide, and alginate	Technical University Hamburg, Germany	-
PAO1 ΔfliC wt algC ΔpilA	Lack of flagella; lack of type IV pili		-
PAO1 ΔfliC wt algC wt pilA	Lack of flagella		-/+
PAO1 wt fliC wt algC wt pilA	Wild type		+

+ activity, bacterial lysis. - no activity.

6.2.6. Host range analysis

The lytic activity of PA5oct was examined on Pirnay's *P. aeruginosa* panel (Pirnay *et al.*, 2002), excluding "lysis from without" spots. Phage activity spectrum was tested against 57 clinical isolates from Military Hospital Neder-Over-Heembeek, Brussels, Belgium (Table S 4.1). PA5oct was able to infect 26% strains, whereas representatives of the *Luz7virus* group (LUZ7 32%), *Lit1virus* group (LIT1 17%), *Luz24virus* group (LUZ24 28%), *Phikmvvirus* group (LUZ19 40%, LKD16 39%, LKA1 5%, phiKMV 37%), *Pbunavirus* group (LBL3 46%, LMA2 21%, LSL4 17%, KT28 28%, KTN6 42%) and *Phikzvirus* group (KTN4 35%, φKZ 37%). Phage narrow host range was also confirmed by de Soyza *et al.* (2013) on 43 clinical *P. aeruginosa* strains from COST LMG international reference panel. The PA5oct phage was able to infect 24% isolates, compared to representatives of the *Luz7virus* group (LUZ7 42%), the *Phikmvvirus* group

(LUZ19 44%), the *Pbunavirus* group (LBL3 39%, KT28 28%, KTN6 42%) and the *Phikzvirus* genus (φKZ 46%, KTN4 33%).

6.2.7. Phage influence on biofilm

To initially evaluate PA5oct biofilm eradication ability, 24 h and 48 h old *P. aeruginosa* PAO1 biofilms were grown on 8-well slide microscopy chambers and further infected with PA5oct suspension (6.5 x 10^{10} pfu/ml), and incubated for 24 h at 37°C. Biofilm treated with LB broth served as a negative control and treatment with a high dose of gentamicin (400 μg/ml) as a positive control. The MIC of gentamicin for tested strain was 0.5 μg/ml. Biofilms were stained with LIVE/DEAD *Bac*Light Bacterial Viability Kit and visualized by confocal laser scanning microscopy (Fig. 6.7). We observe that in both 24 h and 48 h old biofilms, PA5oct and gentamicin treatments altered the structure of biofilm in comparison to the control. Biofilm negative controls presented a dense structure, with a thickness that increased over the time. However, after phage or gentamicin treatments, the biofilm structure was less dense with clear areas of biofilm disruption. This study shows the initial assessment of PA5oct antibiofilm activity, which was further evaluated using more qualitative and quantitative techniques by Drulis-Kawa group (Department of Pathogen Biology and Immunology, Institute of Genetics and Microbiology, University of Wrocław, Poland) and collaborators (Department of Microbiology, Institute of Biology and Department of Molecular Physics, Institute of Physics, The Jan Kochanowski University in Kielce, Poland and Zeta Instruments Co., San Jose, USA). Results are are currently in preparation for publication.

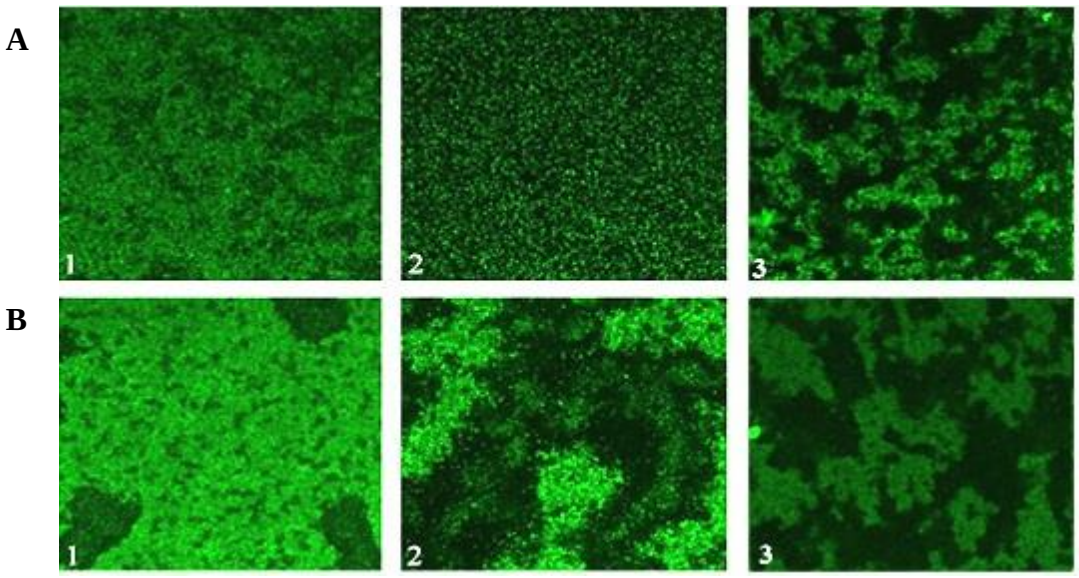


Figure 6.7: Antibiofilm activity of phage PA5oct tested on *P. aeruginosa* PAO1 biofilms. (A) 24 h and (B) 48 h old biofilm. 1) negative control: biofilm treated with LB broth, 2) positive control: biofilm treated

with gentamicin (400 μg/ml), 3) PA5oct treatment (6.5 x 10^{10} pfu/ml). Biofilms were grown at 37°C, 5% CO_2, stained with the LIVE/DEAD *Bac*Light stain (LIVE/DEAD *Bac*Light Bacterial Viability Kit) and visualized by confocal laser scanning microscopy (Zeiss LSM 510 Meta 40× objective, Jena, Germany). Live bacteria are stained green.

6.2.8. Antibacterial efficacy of PA5oct phage analyzed in gentamicin exclusion assay on Airway Surface Liquid infection model

The *in vitro* antibacterial activity of PA5oct was assessed using the gentamicin exclusion assay on ASL. The experimental set up was identical to the experiment with phage KTN4, section 5.1.8., including NuLi-1 and CuFi-1 bronchial epithelium cell lines (Zabner *et al.,* 2003), and three *P. aeruginosa* strains (PAO1, nonCF0038 and CF708) (Olszak *et al.,* 2015). The epithelial cells viability controls were established, as was shown in section 5.1.8. No toxicity influence of cell lines was observed for phage and bacterial samples.

First, both epithelial cell lines were infected with selected strains for 3 h and colony count showed, that all *P. aeruginosa* strains could efficiently propagated in both ASLs (10^7 - 10^9 cfu/ml), as it was previously presented in section 5.1.8. Next, the capacity of phage PA5oct to eradicate the extracellular bacterial load was evaluated. CFU counts of *P. aeruginosa* were significantly ($p < 0.05$) reduced for the normal NuLi-1 and CF CuFi-1 epithelia cells. In the case of NuLi-1 a 4.5 log, 6.5 log and 3 log decrease was observed for PAO1, nonCF0038 and CF708, respectively (Fig. 6.8 A). The phage treatment of the CuFi-1 epithelia infection was also very effective ($p < 0.05$) giving 5 log, 2.5 log and 5 log reductions in CFU of PAO1, nonCF0038 and CF708, respectively (Fig. 6.8 B). PAO1 and CF708 was significantly more susceptible ($p < 0.05$) to phage treatment in CuFi-1 cells compared to NuLi-1 cells, contrary to the nonCF0038 isolate. It was confirmed that the PA5oct phage could freely diffuse and gain access to the bacterial hosts in both ASL models, but the final result of the treatment was strongly dependent on the strain features, as it was previously notified in section 5.1.8. (Danis-Włodarczyk *et al.*, 2016; Worlitzsch *et al.,* 2002). In Worlitzsch *et al.* (2002) study *P. aeruginosa* was not interacting with the CF epithelium directly, but was rather found trapped in mucus plugs formed in the airways.

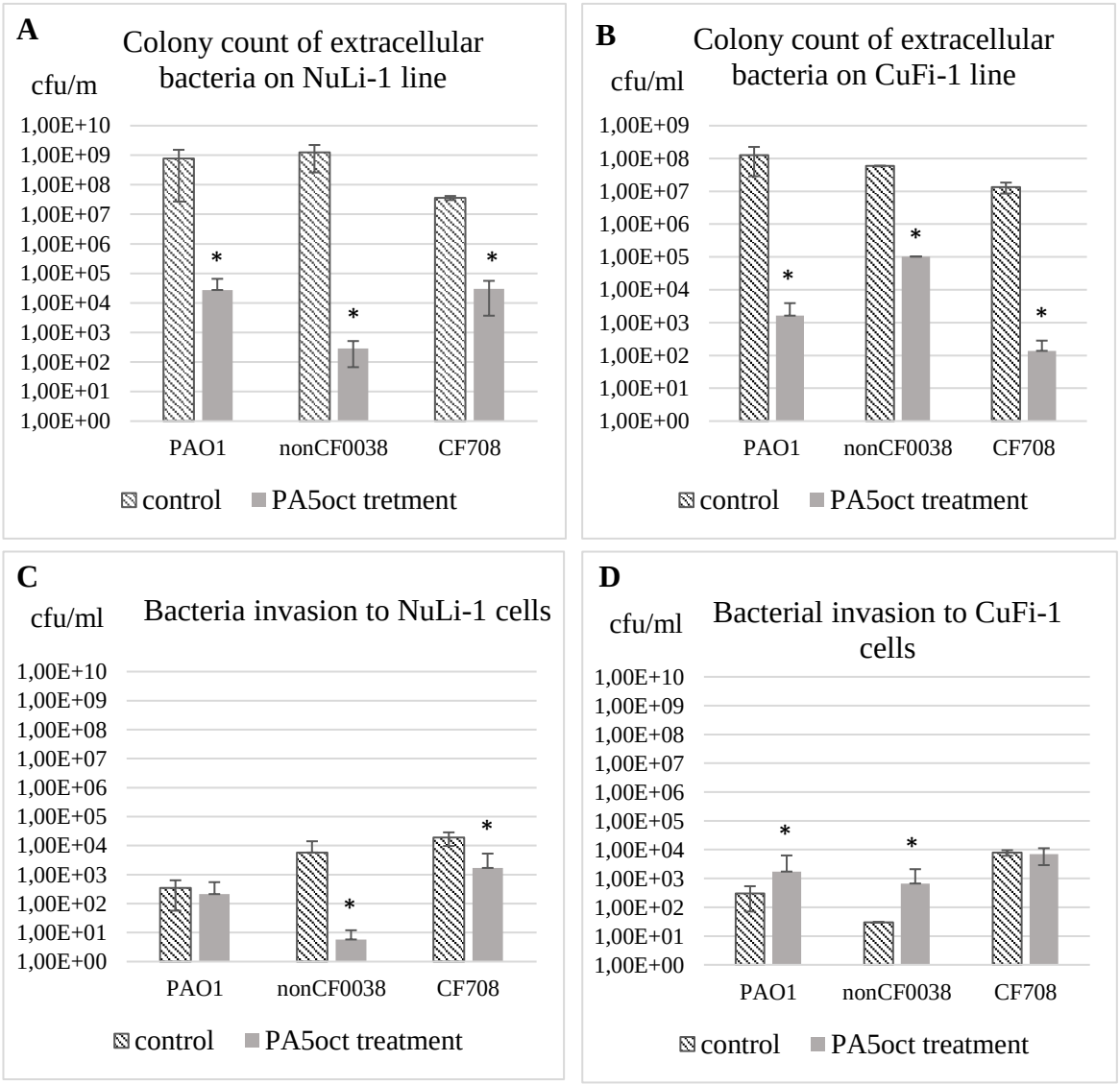


Figure 6.8: Phage PA5oct treatment of *P. aeruginosa* infected NuLi-1 and CuFi-1 epithelial cells. (**A and B**) Colony count of bacteria collected from apical wash; (**C and D**) colony count of bacteria internalized in epithelial cells. The error bars indicate the standard deviation. The results are presented as the means ± SD. Statistical analysis was made by the ANOVA test (denoted p-values < 0.05). (*) p-values < 0.05.

In the final step of the experiment, the ability of *P. aeruginosa* strains to invade into epithelial cells was assessed, as it was presented in section 5.1.8. and the influence of PA5oct phage treatment on the number of invaded bacteria (Fig. 6.8 C and D). The CF708 strain was significantly ($p < 0.05$) more capable to invade the epithelial cells in comparison to PAO1 strain in NuLi-1 line and in compare to both nonCF strains in case of CuFi-1 line. When PA5oct treatment was applied, in the NuLi-1 cell line we observe significant reduction in CFU counts ($p < 0.05$) by 3 logs and 1 log for nonCF0038 and CF708 isolates, respectively (Fig. 6.8 C). Moreover, treatment of nonCF0038 strain was the most efficient ($p < 0.05$). In contrast, no influence on internalization was observed in CuFi-1 line (Fig. 6.8 D). Phage treatment was not

effective and we observe a significant ($p < 0.05$) increase of invasion by 0.5 log and 1 log for PAO1 and nonCF0038, respectively. In case of CF708 there is no significant difference between untreated and treated samples.

6.3. Genome analysis

6.3.1. Genome properties

Phage PA5oct has a giant linear A+T-rich (33.3% GC) double-stranded DNA genome (287,182 bp), making it the third largest sequenced *Pseudomonas* phage genome and the eighth jumbo genome among *Myoviridae* phages (Table 6.1). Based on DNA sequencing 397 genes were predicted with the use of GeneMark S, GeneMark.hmm, OrfFinder and manual inspection (Fig. 6.9).

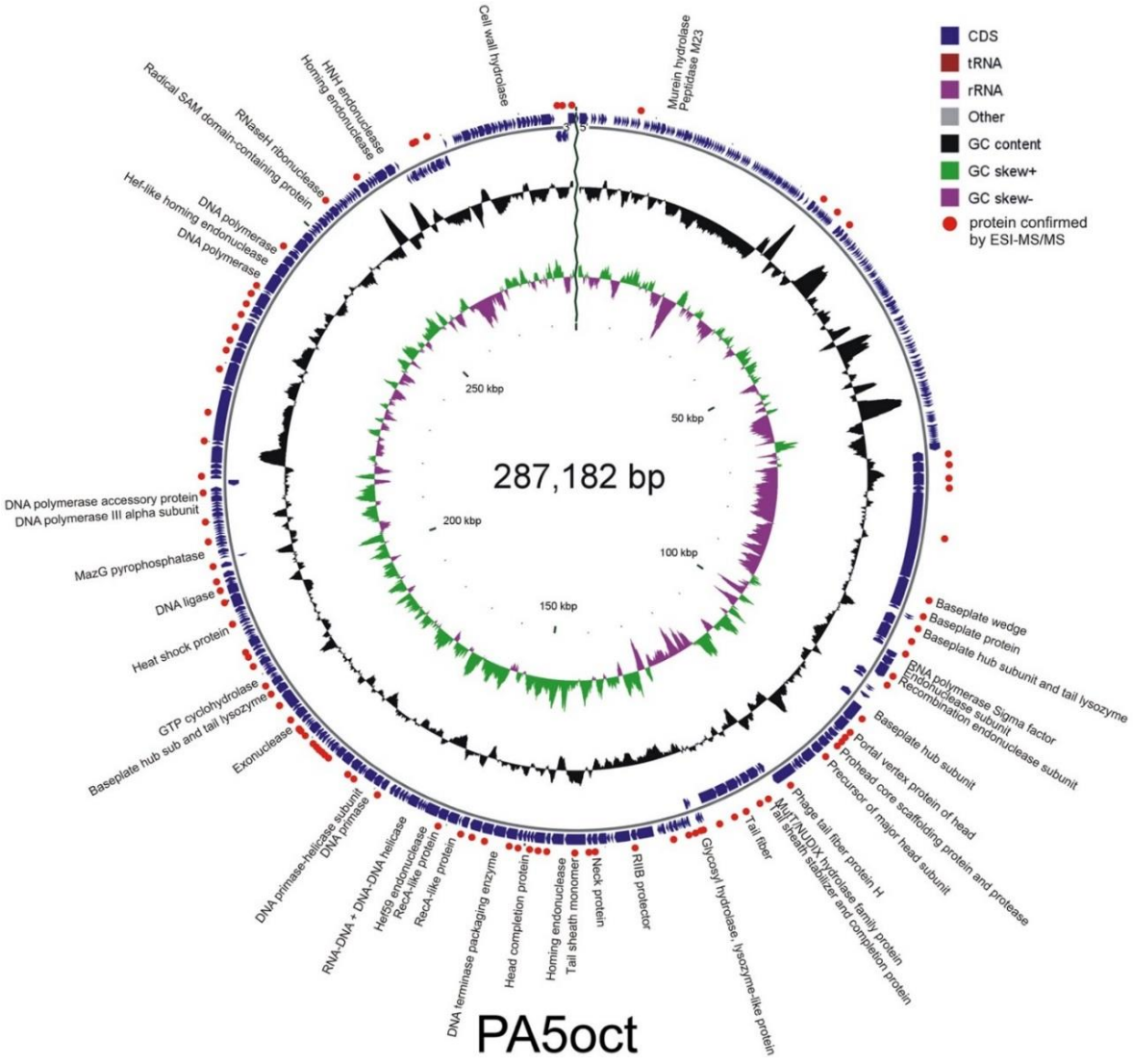


Figure 6.9: Circular representation of the PA5oct genome. Blue arrows represent predicted ORFs on the Watson and Crick strands. The inner black circle represents the GC content and the green-violet represents GC skew. Structural proteins confirmed by ESI-MS/MS analysis are marked with red dots.

RNA sequencing allowed for identification of 65 additional ORFs (+14.1%) and confirm presence of previously predicted genes (Fig. 6.10). Alternative start codons were proposed for nine genes (encoding gp37, gp169, gp283, gp294, gp330, gp333, gp341, gp372, gp444). In total, 462 ORF's (open reading frames) and ten tRNAs [Met (CAT), Leu (TAA), Asn (GTT), Met (CAT), Leu (TAA), Arg (TCT), Pro (TGG), Gly (TCC), Ser (TGA) and Ser (GCT)] were identified (Table S 6.2). This result can be comparable to the smallest known bacterium, *Mycoplasma genitalium* G37, which has 470 genes (Fraser *et al.*, 1995).

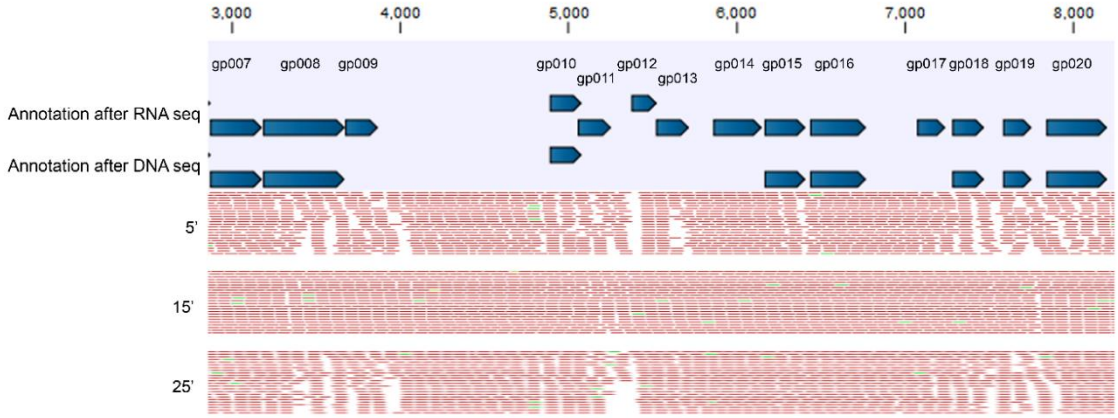


Figure 6.10: Comparison of PA5oct genome annotation in the region between 3 and 8 kb based on DNA and RNA-seq. The reads on a Watson strand are marked red for samples taken 5, 15 and 25 min post infection.

According to the orientation of transcription, ORFs are organized into operons and most of them are located on the positive strand (Fig. 6.11). The PA5oct protein sizes range from 33 to 3858 amino acid residues, within 8.87% (41 ORF's) is ≤ 50 amino acids, 47.83% (221 ORF's) is ≤ 100 amino acids, 45.45% (210 ORF's) is > 100 and <500 amino acids, and 6.49% (30 ORF's) ≥ 500 amino acids. Based on bioinformatics and ESI-MS/MS analysis function was assigned to 117 genes (25.32% of all proteins), including 73 structural proteins (62.39% of proteins with appointed function and 15.80% of all proteins). Importantly, none of them were related to integration proteins (integrase, transposase) or antimicrobial resistance, which allows to consider this phage as a potential antimicrobial.

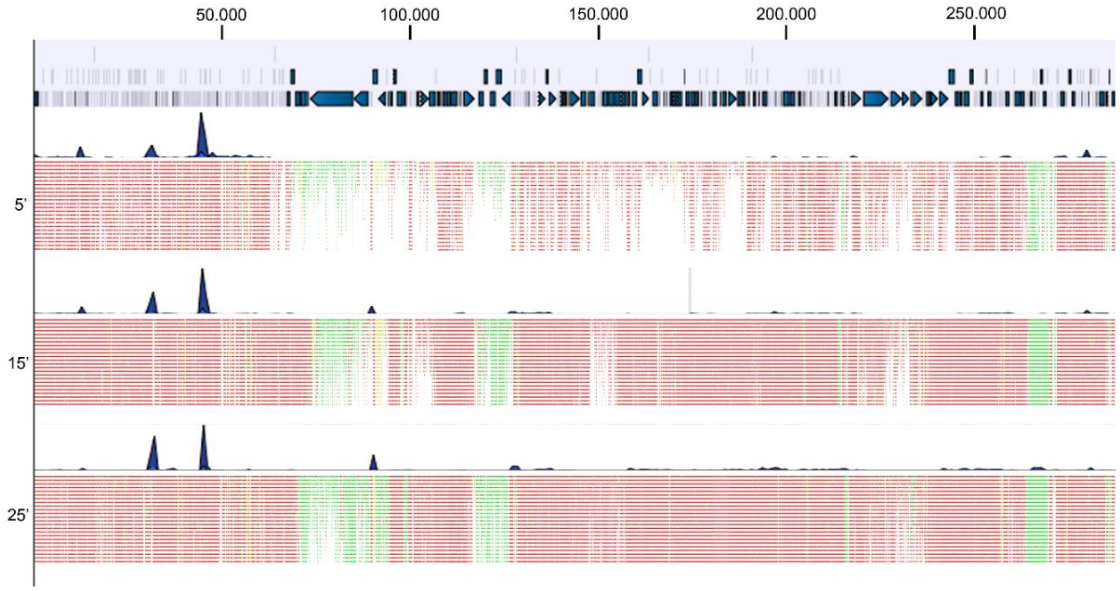


Figure 6.11: RNA-seq analysis of the PA5oct transcriptome after *P. aeruginosa* PAO1 infection. Genome-wide overview of reads mapped to the Watson (red) and Crick (green) strand of PA5oct genome for samples taken 5, 15 and 25 min post infection. The blue graph on top present the relative abundance of each position of the genome normalized by the total amount of sequencing reads in each sample.

Furthermore, 39 promoters with highly conserved, AT-rich intergenic motifs (5'-TATAATA-3') and (5'-TTGAC-3') were identified (Fig. 6.12).

		-35	-10	
ORF 1	+	TTGACACAGTTACTGTTTAATTG	TATTAT	5-33
ORF 2	+	TTAGCACAAAAATATTGTTTTATGC	TATAAT	362-392
ORF 10	+	TGTTGCAATAGCTTAAAAGTTTTAG	TATAAT	4103-4133
ORF 14	+	TTGACCAAACGCACAAAACAGTA	TATAAT	5737-5765
ORF 20	+	TTGACACAAGTATAATTATTTCG	TATAAT	7779-7807
ORF 40	+	TTGACTATCCATTCGTGTTGGGTTAATTTAAC		15190-15221
ORF 100	+	TTGACGCAATGCAAACTAGGCTC	TACACT	32126-32154
ORF 114	+	TTGACATTTGATGTGTATACTAG	TATAAT	35886-35914
ORF 116	+	TTGTGTTTTCTATTTTGAAGTGC	TACATT	38012-38040
ORF 117	+	TTGCTATACGTACTAATACGTTG	TATACT	38642-38670
ORF 127	+	TTGCTATTATGTGTTTATAGAAG	TTTAAT	41992-42020
ORF 132	+	TTGTGCAATGCATTTGTTGGTGT	TATATT	43858-43886
ORF 148	+	TGTTGACGCATATAGCAATATGCGC	AATAAT	48302-48332
ORF 181	+	TTGACACCTATATGTA GAATATAG	CACAAT	54155-54184
ORF 165	-	TGCATTAGCACACACATAGCG	TATACT	55568-55541
ORF 168	+	TTGACAAAAAATAATAATTAAAAA	TATAAT	56447-56476
ORF 171	+	TTGTACATGGTGGCCAGATCTGCA	TATACT	57427-57456
ORF 173	+	TTGACATGTACATAAGATCATAT	TATGCT	58645-58673
ORF 178	+	TTGACGCAGACTAGCATCGGACA	CATAAT	59953-59981
ORF 184	+	TTGACGTAGCCGAGCGAAGGGCG	CATAAT	61741-61769
ORF 212	+	TTGACGGATAACAGTAAAATTAG	TATACT	89385-89413
ORF 217	-	TTGTGCATCCATGCTTTTTATTAT	TATTTT	94539-94510
ORF 218	+	TTGACTTTGAAAATATTTTGTTG	TAAAAT	94486-94514
ORF 234	+	TTGACTAGAAAAAAATAATTTGT	TAGTAT	110863-110891
ORF 258	+	TTGACGTTATATTTTAATTGTGT	TTTACT	131770-131798
ORF 263	+	TTGACGAAATACATTGAATATTT	TAAAAT	133678-133706
ORF 273	+	TTGACTATAAATTATATCTATAG	TATACT	145097-145125
ORF 303	+	TTGACGCAATAACGGGATTTTTGC	TATTAT	174273-174302
ORF 316	+	TTGACATATTTATAGCCTTATAT	TATAAT	182009-182037
ORF 329	+	TTGACACGGAAGAAAACGTATAG	TATACT	192155-192183
ORF 333	+	TTGCAGAAATCGTAGGAATAAACTC	TATTAT	195007-195037
ORF 336	+	TTGACGACACTGTTTAAACATA	TATAAT	196341-196368
ORF 364	+	TTGACAAGTGCCCTTTTTTTGCCA	TATAAT	210279-210308
ORF 377	+	TTGACAACAAGTTTCAGTAACG	TACTAAT	219682-219710
ORF 388	+	TTGACTAACTCCGTTTACACTAG	TATAAT	240017-240045
ORF 392	+	TTGCATTATGAAAATATTTAATG	TATAAT	245650-245678
ORF 431	-	TTGACTATTGAATATACATATTT	TATAGT	269810-269782
ORF 432	+	TTGACATTGGTATTGTCTTTTAG	TAAACT	269845-269873
ORF 438	+	TTGACTCATTTGTTGGATAGCAC	TATACT	272443-272471

Figure 6.12: Alignments of PA5oct promoters. The conserved AT-rich intergenic motifs: -10 box (5'- TATAATA - 3') and -35 box (5' - TTGAC - 3') are marked with pink colour. The corresponding sequence logos from MEME/MAST are depicted below the alignments. On the left site are listed ORFs in front of which promoters are located, on the right side is sequence locus in the phage genome.

Using ARNOLD software and RNA seq analysis, 60 putative factor-independent terminators were predicted with conserved stem loops (Fig. 6.13).

```
ORF 001 + 313-353           Both + AAATCAATAAAAAGGGCTTAAAGCCCTTTTTATTTGTCTGC              - 8.80
ORF 028 + 11539-11585       Both + TATTAATAAAACATTGCCCGCATTAGCGGGCAATGTTTTATTGTATT        -13.70
ORF 032 + 12841-12881       Both + AAATATAAAAAAGGGCACATTGTGCCCTTTTTTAATAGTAA              -10.60
ORF 032 + 12910-12956   Rnamotif + TCTATAATATACACTGTGCATTCTAATACACAGTGTATTGTTGCGAG        - 8.50
ORF 098 + 31595-31638       Both + TTATATCAGAAAGGCAGACTTATGTTTGCCTTTCTTTTTGAGCA          -11.30
ORF 100 + 32950-32994       Both + ACTGCGCCGAAAGGGGAGCAAAATGCTCCCCTTTCTTCATGACCA         -16.00
ORF 115 + 37920-37966       Both + ATGCATATTTAGAAAGGGAATTATAATTCCCTTTCTTTTTGGATCTA       -11.40
ORF 128 + 42749-42794       Both + GCTGAGAAAAAGAAGCAGGAATTATTCCTGCTTCTTTTGTATGTTA        -13.90
ORF 131 + 43697-43743       Both + CGTTTGATAATAAAGGGGGAATTATCCCCCTTTaTTTTTCTTTTATA       -14.20
ORF 146 + 47682-47728       Both + TACATTTATTAGGAAAGGGGCTTAAACCCCTTTCCTTTTGATTGTTA       -14.90
ORF 149 + 49254-49297       Both + TTAAATAGAAAGGAAGGGGTTAAACCCCTTCCTTTTATTGTAGG          -16.30
ORF 153 + 51068-51114       Both + CTGCTAACACAGGAAAGGGGCAGTTGCCCCTTTCCTTTTATGTCCTA       -14.90
ORF 155 + 51653-51698       Both + AAACAATAAGAGAAAGGGGCAACTGCCCCTTTCcTTTTATGTCCTA        -14.20
ORF 157 + 52370-52417       Both + CAATTACTAAAGGAAAGGGGCAGTTGCCCCCTTTCCTTTTATCCATTA      -14.10
ORF 158 + 52912-52956       Both + AGCACAAAAAGGAAGGGGCGCAAGCCCTTTTcTTTTGTCTGCAAT        -14.50
ORF 158 + 53056-53096       Both + GCATAACACAAAGGGCTTGACAAGCCCTTTGTGTTATGCGT             - 8.80
ORF 159 + 53613-53652       Both + TAACAAAGGAAAAGGGCAACTGCCCTTTTCCTTATGACCA              - 9.10
ORF 160 + 54037-54077       Both + TAATTAAGAAAAGGGCTATTATAGCCCTTTTCTTATGACCG             - 9.20
ORF 163 + 55234-55269   Rnamotif + TCATAAGAAAAGGGGCATTGCCCCTTTTCTTATAGT                  - 9.00
ORF 164 - 55230-55270      Erpin - GACTATAAGAAAAGGGGCAATGCCCCTTTTCTTATGACTAT             -10.50
ORF 173 + 58936-58981       Both + TAGTGACAAAAGAAAGGGGAATTAATCCCCTTTCTTTTTATTAAGT        -13.80
ORF 180 + 60558-60600      Erpin + CACCAACTAAAGCCCCGCCTCGCGCGGGGCTTTACCGAAGGAA           -15.80
ORF 212 + 89713-89757       Both + AACAGAAGGGAGGGGGTTAAGTAACTAACCCCCTATTTTTATGCC         -14.10
ORF 229 + 106911-106949 Rnamotif + TGTTAGTTTTTGGGGCATAAATGCCCCTTTTTTAATATT              -11.30
ORF 223 + 110824-110869    Erpin + ATACCAAAAAAAGTGTAGTAGAAATACTACACTTTTTTATTGACTA       - 9.20
ORF 237 + 114012-114052     Both + TAAAATAAAAAAGGAGCATAATGCTCCTTTTTTATTGATAA            - 9.10
ORF 264 + 136674-136713     Both + GTACATAAAAAGGGATGCATTGCATCCCTTTTTATTTTAT             -11.30
ORF 266 + 139292-139337 Rnamotif + GCATAAGCAGCTAGGAGACTTAAAACGTCTCCTATTTTTATATATA      -11.50
ORF 271 + 142289-142328     Both + AATAATATAAATAGGGGCAATGCCCCTATTTTATTGAGTG             -11.40
ORF 272 + 145064-145104     Both + TGCAAAATAAAAAGGCACTTAGTGCCTTTTTATTTGACTAT            - 8.20
ORF 274 + 147476-147524     Both + AATAAGAAATCTGCAAGCTATTTTACTTAGCTTGCAgTTATTCTTTTTA    -10.60
ORF 284 + 158413-158460     Both + AACTAAATAGAGAGAGGGGATCGCTGATCCCCTCTCTTTAATATAACC     -16.00
ORF 303 + 174250-174303     Both + TTCTATAAAAATCCCGTTACTGCTTGACGCAATAACGGGAtTTTTGCTATTATA - 9.52
ORF 309 + 178457-178502     Both + AAATAAAACTTGAAAGGGCTTATTAAGCCCTTTCTTGTTCTTAGTG       -13.30
ORF 319 + 186683-186723     Both + ATAATATAAAAAGGGAACAATGTTCCCTTTATTACACACAA            - 8.40
ORF 335 + 196295-196334     Both + TACAAATAAAAAGCTGGTATTCCAGCTTTTTATTTTGTTA             - 8.80
ORF 336 + 196709-196752     Both + ACAGCAAAAACGGGGGTAGACACCTGCCCCCTTCCCTATTATAA         -16.20
ORF 340 + 198228-198269     Both + AGTAAAAACAATCCCCGCAATTGCGGGGATTGTTTTATAAGG           -14.70
ORF 345 + 202786-202826     Both + GTGTAAAATAAAAGGGCGCTGAGCCCTTTTATTTTATCATA            - 9.50
ORF 346 + 203652-203695    Erpin + TAAGATATAAAAATGCCCCATACGGGGCATTTTTATCTATCCAT         -13.20
ORF 349 + 204841-204879     Both + TGATAATTTAAGGGGACTTCGGTCCCCTTTTTTTGCTTG              -15.10
ORF 349 + 204940-204980     Both - ATTACTAAAAACCCCGCATAATGCGGGGTTTTTAATGCATT            -12.70
ORF 363 + 210262-210303     Both + TAGTTGTAAAAGGGCACTTGACAAGTGCCCTTTTTTTGCCAT           -12.50
ORF 371 + 214401-214442    Erpin + ACAAAAATAAAACCCTAACTAGTTAGGGTTTTATTTTGGGTT          - 8.90
ORF 372 - 214399-214440    Erpin - CCCAAAATAAAACCCTAACTAGTTAGGGTTTTATTTTTGTTA          - 8.90
ORF 372 - 214450-214489     Both - AATTAAATAAAAGGGGCATATGCCCCTTTTATTTTGTTTA            -11.50
ORF 372 + 214452-214492     Both + AACAAAATAAAAGGGGCATATGCCCCTtTTATTTAATTAAT           -11.50
ORF 374 + 216683-216727    Erpin + ATAATCAAAAAGCACTTATTTTCAATAAGTGCTTTAACTATCATG       - 9.20
ORF 376 + 219648-219688     Both + ACTCCTAAAAATAGCCCCGAACGGGGCTATTTTTTTGACAA           -13.50
ORF 379 + 227301-227340     Both + TAGTTAATAAAAGGGCTTTTAAGCCCTTTTATTATATTTT            - 8.80
ORF 383 + 235858-235897     Both + AAATTAAATAAAGGGCCCTAGGGCCCTTTATTTATGTATA            -12.60
ORF 384 + 236506-236548     Both + GGAATTTGTTAGCTGACCCGTTTGGGTCAGCTCTTTTATTGGA         -15.70
ORF 413 + 259005-259045     Both + GGCAAAATAAAAGGGGCTTAAAGCCCCTTTTATTTTATATA           -11.20
ORF 418 + 263080-263118     Both + AAAACAAAAAAGGGGCTTATAAGCCCCTTTTTTGTTGAC             -11.50
ORF 419 + 263641-263686    Erpin - TTGATTAAAATAAAAGGGCTTAACGCCCTTTTaTTTTGTAAATTAG     -11.00
ORF 421 - 264135-264175     Both - AACTAACAAAACAGCCCTAGCTAGGGCTGTTTTTTTGGGCA           -12.30
ORF 428 - 267126-267166     Both - GTAATGCTAAAGGGGAATTTAATTCCCCTTTTAATTTATAT           - 9.50
ORF 432 + 270853-270894     Both + AGTCATTAAAAACCCTACATTTGTAGGGTTTTTATTTGATTA          -10.50
ORF 444 + 277206-277244 Rnamotif + TAACACAATAACGCGCTTTACTGCGCGTTAATTTGCGCG            - 9.30
ORF 450 + 279512-279555     Both + TTTATATAAAATAAGGGACTTTATGTCCCTTATTTTATATAAGA        -12.30
```

Figure 6.13: Phage PA5oct predicted terminators with palindromes marked blue. From the left site: list of ORFs after which terminators are located, strand, location in phage genome, motif of terminator, terminator sequence with conserved stem-loop structures marked blue, Δ*G* (the Gibbs free energy of stem-loop formation in kcal/mole).

6.3.2. Transcriptomics analysis with RNA-seq

Reads originating from the phage and the host at each stage of infection were mapped to the phage and host genomes respectively, revealing that PA5oct progressively dominates host transcription. Indeed PA5oct transcripts represented 21% of total transcripts within 5 minutes and eventually proceed to 69% then 92% by middle and late infection, respectively (Fig. 6.14). The strand specificity of the results showed that early transcription is almost exclusively performed on the Watson strand, in sense with what can now be defined as early coding sequences. Specifically, gp32, gp33, gp99 and gp450 (uncharacterized proteins) can now be defined as early genes, which are classically understood to be typically involved in the transition from host to phage metabolism. PA5oct also appears to have a distinct middle phase of transcription captured at 15 min centering around the region between 70 kb and 200 kb with gene features predicted to be responsible for structural proteins as well as DNA metabolism and replication. Outside of this region, gp99 and gp213 (uncharacterized proteins) are particularly strongly expressed in middle infection, which continues into late infection, where gp99 is the second most strongly expressed protein. In the late phase of infection, gp250, a glycosyl hydrolase lysozyme-like protein, is more strongly expressed compared to the early and middle phase, which could suggest that it is required for destabilization of cell wall and phage progeny release.

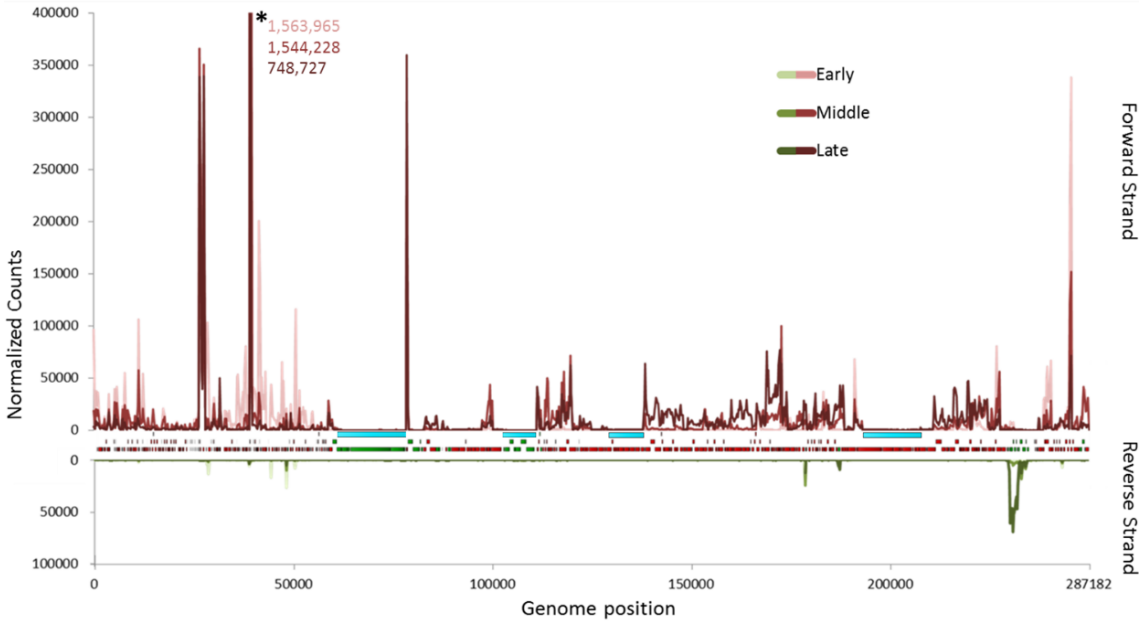


Figure 6.14: RNA-seq analysis of the PA5oct transcriptome after *P. aeruginosa* PAO1 infection. Genome-wide overview of reads mapped to the Watson (red) and Crick (green) strands of PA5oct genome for samples taken 5, 15 and 25 min post infection. Four barely transcribed regions are highlighted with blue bars. Each of the three samples were normalized to each other by the total count of phage reads.

Additionally, five non-coding RNAs which lack similarity to any known DNA sequence in GenBank could be described (Table 6.3). Notably, ncRNA003 produces 44%, 46%, and then 29% of phage transcripts over the course of the infection process, even though this feature comprises only 235 bp of the 287 kb genome.

Table 6.3 Identified non-coding RNA species.

Name	Location	Strand
ncRNA001	4113..4883	+
ncRNA002	31367..31616	+
ncRNA003	44676..45001	+
ncRNA004	126878..127154	+
ncRNA005	269893..271298	-

Remarkably, four regions, including two of the three major regions transcribed on the Crick strand, were found to be only barely transcribed in spite of significant sequencing depth. This could potentially reflect either recently acquired morons that have not yet been adapted to the transcriptional strategy of PA5oct, or perhaps these regions could require host transcription factors not found in PAO1 for optimum expression.

6.3.3. ESI-MS/MS proteome analysis

ESI-MS/MS analysis of gel-separated phage particle proteins led to the experimental identification of 93 virion-associated gene products, including seven virion-unrelated enzymes, 13 virion-associated proteins and 73 structural gene products, with sequence coverages ranging between 2.9% to 98.0% (Fig 6.15, Table S 6.3). The PA5oct particle proteins were compared to their homologues from *Cronobacter* phage vB_CsaM_GAP32 (JN882285.1), *Escherichia* phage PBECO 4 (NC_027364.1), *Escherichia* phage 121Q (NC_025447.1), *Enterobacteria* phage vB_KleM_RaK2 (NC_019526.1) and *Klebsiella* phage K64-1 (NC_027399.1) (BLASTP). The highest amino acids similarity was found with *Cronobacter* phage vB_CsaM_GAP32 ranging from 28% to 68% for 23 proteins, *Escherichia* phage 121Q ranging from 21% to 66% also for 23 proteins and *Enterobacteria* phage vB_KleM_RaK2 ranging from 22% to 66% for 22 proteins.

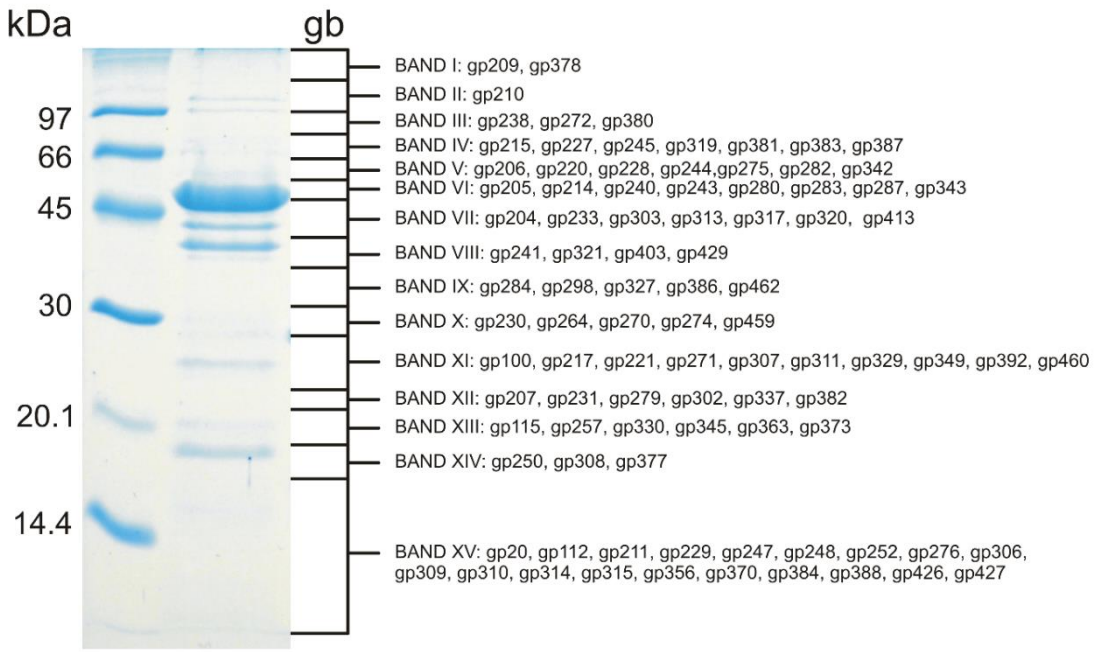


Figure 6.15: The SDS-PAGE pattern of PA5oct structural proteome against LMW Ladder (Thermo Scientific) in first line. The corresponding molecular weight is mentioned left. The numbered fractions on the right, correspond to gel slices analyzed individually by ESI-MS/MS. The proteins are mentioned in the slice in which they were most abundantly present.

6.3.4. Functional annotation of PA5oct genome

Capsid

ESI-MS/MS analysis identified three structural head related proteins, gp228 (portal vertex protein of head), gp231 (prohead core scaffolding protein and protease) and gp233 (precursor of major head subunit). Additional protein, gp277, a head completion protein, was annotated based on homology with the *Enterobacteria* phage vB_KleM_RaK2 head completion protein (e-value 1E-32).

Tail apparatus

Neck protein, gp270 and four tail proteins, gp238 (phage tail fiber protein H), gp240 (tail sheath stabilizer and completion protein), gp243 (tail fiber protein) and gp272 (tail sheath monomer) as well as three baseplate proteins, gp210 (baseplate wedge), gp211 (baseplate protein) and gp214 (baseplate hub subunit and tail lysozyme) were identified during ESI-MS/MS analysis. Gp238, putative phage tail fiber protein H, has on C-terminal putative endo-N-acetylneuraminidase region (868 - 1008 amino acids position), which has possibility to hydrolase (2→8)-alpha-sialosyl linkages in oligo- or poly(sialic) acids, activity associated with the tail spikes and EPS depolymerases (Kwiatkowski *et al.*, 1982).

Peptidoglycan degrading enzymes

Several genes associated to the peptidoglycan layer degradation were identified: gp45, gp46, gp214, gp250, gp321 and gp447 (BLASTP, HHpred, HMMER, InterPro, Phyre2). Gp45 is a predicted murein hydrolase, peptidase M23 with a 103 amino acids long domain (30 - 133 aa) and gp46 is a putative peptidase M, both capable of cleaving bonds in polymeric peptidoglycan. Gp214 is a putative baseplate hub subunit with tail-associated lysozyme (ESI-MS/MS hit), homologues to gp05 of *E. coli* bacteriophage T4, which has a needle-like structure attached to the end of a tube through which the DNA passes on its way out of the head and into the host. The gp05 needle punctures the outer cell membrane and then digests the peptidoglycan cell wall in the periplasmic space through hydrolysis of (1→4)-beta-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in a peptidoglycan and between N-acetyl-D-glucosamine residues in chitodextrins (Kanamaru *at al.*, 2005). Gp250 was assigned as glycosyl hydrolase, lysozyme-like protein with a 159 amino acids long domain (11 - 170 aa) and it can possibly assists in the hydrolysis of glycosidic bonds between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. Gp321 is a putative baseplate hub and tail lysozyme, a muramoyl-pentapeptidase (ESI-MS/MS hit). It contains a PAAR motif (61 - 97 aa) and a muramyol-pentapeptide carboxypeptidase hydrolase (180 - 320 aa), which cleaves a peptide bond at the carboxy-terminal (C-terminal). Finally, gp447 a cell wall hydrolase domain protein, SleB (domain between 80 - 186 aa), which encodes N-acetylmuramyl-L-alanine amidase and cleaves the link between N-acetylmuramoyl residues and L-amino acid residues in cell-wall glycopeptides.

DNA metabolism and genome replication

PA5oct encodes proteins required for the formation of the replisome, comprising a helicase (gp292), several endonucleases (gp219, gp220, gp228, gp273, gp390, gp403, gp417, gp418), exonuclease (gp316), RNA polymerase sigma factor (gp218), DNA polymerases (gp346, gp366, gp369, gp389, gp391), DNA primase (gp299, gp300), DNA ligase (gp343) and DNA terminase (gp281). Also RecA- like proteins (gp285, gp287) were annotated, that play a role in homologous recombination, DNA repair and induction of the SOS response. Gp239 encodes a putative MutT/NUDIX hydrolase family protein involved in diverse metabolic pathways. Moreover, a GTP cyclohydrolase

(gp329) was identified, responsible for the hydrolysis of guanosine triphosphate (GTP) to form 7,8-dihydroneopterin triphosphate (7,8-DHNP-3'-TP, 7,8-NH_2-3'-TP) and MazG pyrophosphatase (gp356), which catalyzes the hydrolysis of all eight canonical ribonucleoside triphosphates (NTP) and deoxyribonucleoside triphosphates (dNTP) to their corresponding nucleoside monophosphates [(d)NMP] and PPi, and subsequently hydrolyzes the resultant PPi to Pi (Zhang *et al.*, 2003). Finally gp400 is a radical SAM domain-containing protein, that catalyze diverse reactions, including unusual methylations, isomerization, sulfur insertion, ring formation (Sofia *et al.*, 2001).

6.3.5. Comparative genome analysis and protein-sharing network

To better represent genetic relationships between the jumbo bacteriophages with genomes of more than 200 kbp (Hendrix, 2009) and other sequenced viruses, a gene-sharing network was built. Viruses that do not show significant similarity to 42 jumbo phages were excluded for clarity. The resulting network was composed of 1,355 viral genomes (nodes) belonging to *Myoviridae*, *Siphoviridae*, *Podoviridae*, or uncharacterized phages and 37,955 relationships (edges) between them.

In the network, 21 jumbo phages have been placed into the largest connected component that mostly consists of the double-stranded DNA phages; whereas the remaining 20 phages into three isolated components (Fig. 6.16 A) and a single isolate with no genetic relatives, i.e. *Sphingomonas* phage PAU (data not shown). Of these, the PA5oct, together with 11 jumbo phages belonging to the T4 superfamily and seven taxonomically unclassified myoviruses, fell within a highly interconnected region, i.e., subnetwork (Fig. 6.16 A and Table S 6.4). A subsequent MCL-based network partitioning assigned 168 phage members of this subnetwork into 5 coherent groups of phages (i.e., viral clusters, VCs); these include VC_2 that are rich in the T4-like viruses and VC_8 mainly comprising T4-related cyanomyoviruses as well as VCs_18, 77, and 111 (Fig. 6.16 B and Table S 6.5). Notably, PA5oct and *Escherichia* phages 121Q/PBECO 4, *Klebsiella* phages vB_KleM_RaK2 and K64-1, and *Cronobacter* phage vB_CsaM_GAP32 were grouped into VC_77, with strong connections to each other in terms of genome-genome similarity value (Fig. 6.16 B). The inclusion of phages into the same VC, which is indicative of a group sharing more conserved gene pools (Lima-Mendez *et al.*, 2008), approximately corresponds to genus-level phage group (Lima-Mendez *et al.*, 2008; Roux *et al.*, 2015).

Additional inspection revealed informative connections between jumbo phages and/or other viruses. For example, the *Phicbkvirus* including phiCbK, Magneto, Swift, Karma, Rogue, and CcrColossus as well as phages RSL1, PaBG, and Lu11, which belonged to VCs 87 and 172, respectively (Table S 6.5), formed two isolated components. This discontinuous positioning of two viral groups clearly reflects their lack of genetic relationships to other viruses as evolutionary distinctive braches (Gill *et al.*, 2012, Yamada *et al.*, 2010, Adriaenssens *et al.*, 2012). In addition, the *φKZvirus* including φKZ, φPA3, 201φ2-1, EL, and OBP and eight jumbo phages as well as a smaller phage, i.e., *Vibrio* phage ΦJM-2012 (Jang *et al.*, 2013), were found in an individual component (Fig. 6.16 A). Interestingly, however, phage members in this φKZ-related component were assigned into two VCs 44 and 150 (Table S 6.5), possibly suggesting that the diversity of this group would be much larger than previously known (Jang *et al.*, 2013). Further, in VCs 43 and 184 (Table S 6.4), the *Bacillus* phages G and 0305φ8-36 shared proteins with only smaller phages having genomes of approximately 126 to 168 kbps (data not shown). This observation may support the hypothesis that these giant phages could be derivatives of smaller phages (Hendrix RW. 2009).

Here, a mathematical modeling of a gene-sharing network showed genetic relationships of the giant phages across 1,999 viral genomes at the gene content level. Of 42 jumbo phages, 20 phages fell into three genetically distinct groups that had no recognizable similarity to other viral genomes, whereas nearly 90% of the remaining phages had similarity to the *Tevenvirinae* phages or their relatives. In particular, network positioning and decomposition suggested that PA5oct appears to be most closely related to a recently emerged new subfamily including phages vB_CsaM_GAP32, 121Q, PBECO 4, vB_KleM-RaK2, and K64-1 as distant relatives of the T4 superfamily (Abbasifar *et al.*, 2014). Collectively, a network-based approach provided more comprehensive views for the diversity of giant phages in various viral lineages.

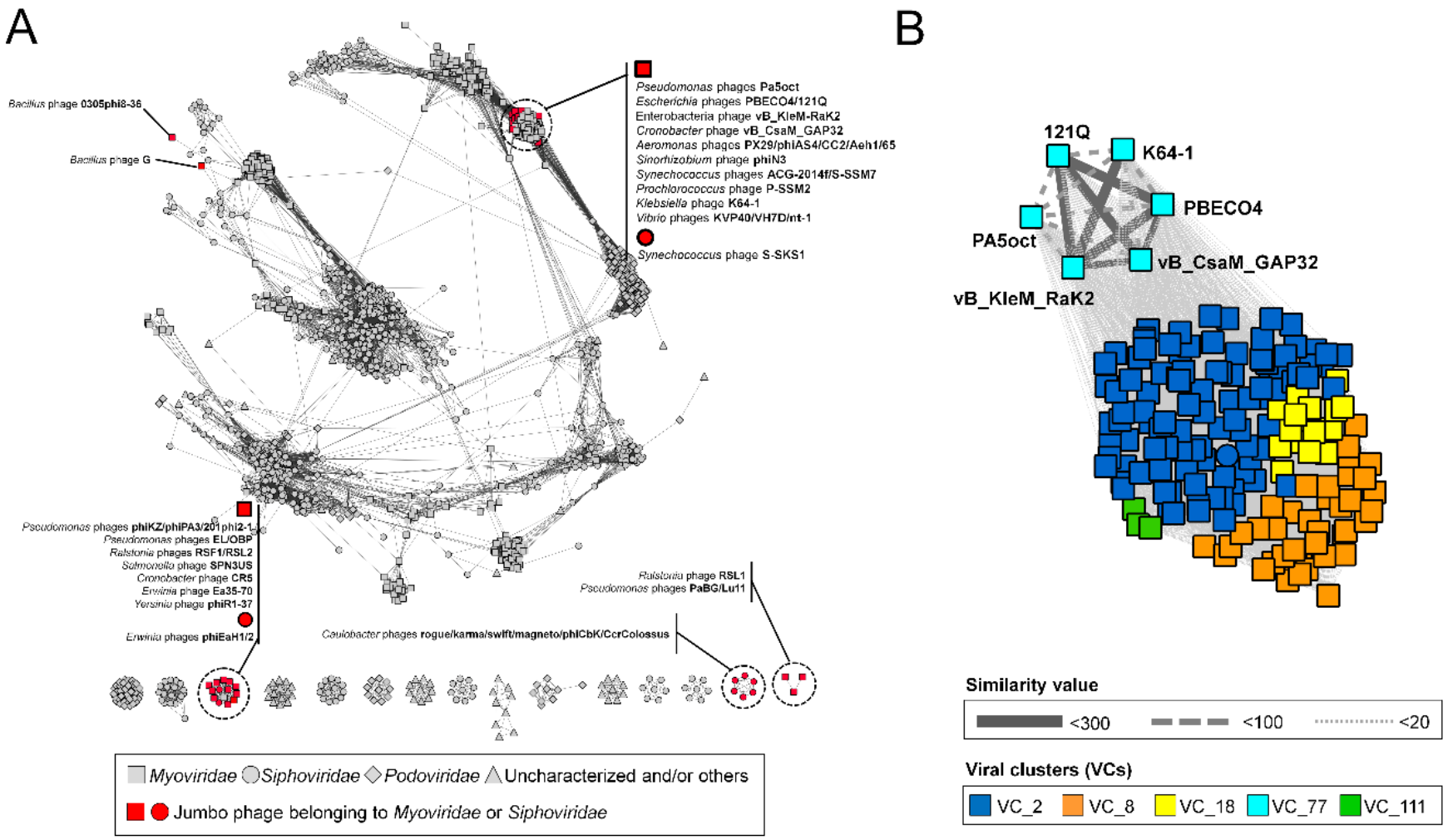


Figure 6.16: Protein-sharing network for PA5oct and jumbo phages. (A) A network representation was produced using the edge-weighted spring embedded layout of Cytoscape version 3.1.1. Each node is depicted as a different shape, representing bacteriophages belonging to *Myoviridae* (rectangle), *Podoviridae* (diamond), *Siphoviridae* (circle), or uncharacterized phages (triangle). Selected jumbo phages are shown in red. Edges between two nodes indicate their statistically weighted pairwise similarities with similarity scores of ≥ 2. (B) An enlarged view of the subnetwork comprising PA5oct and its relatives. Edge thickness is proportional to similarity values estimated with the hypergeometric equation (Materials and methods) and viral clusters are represented in the legend boxes.

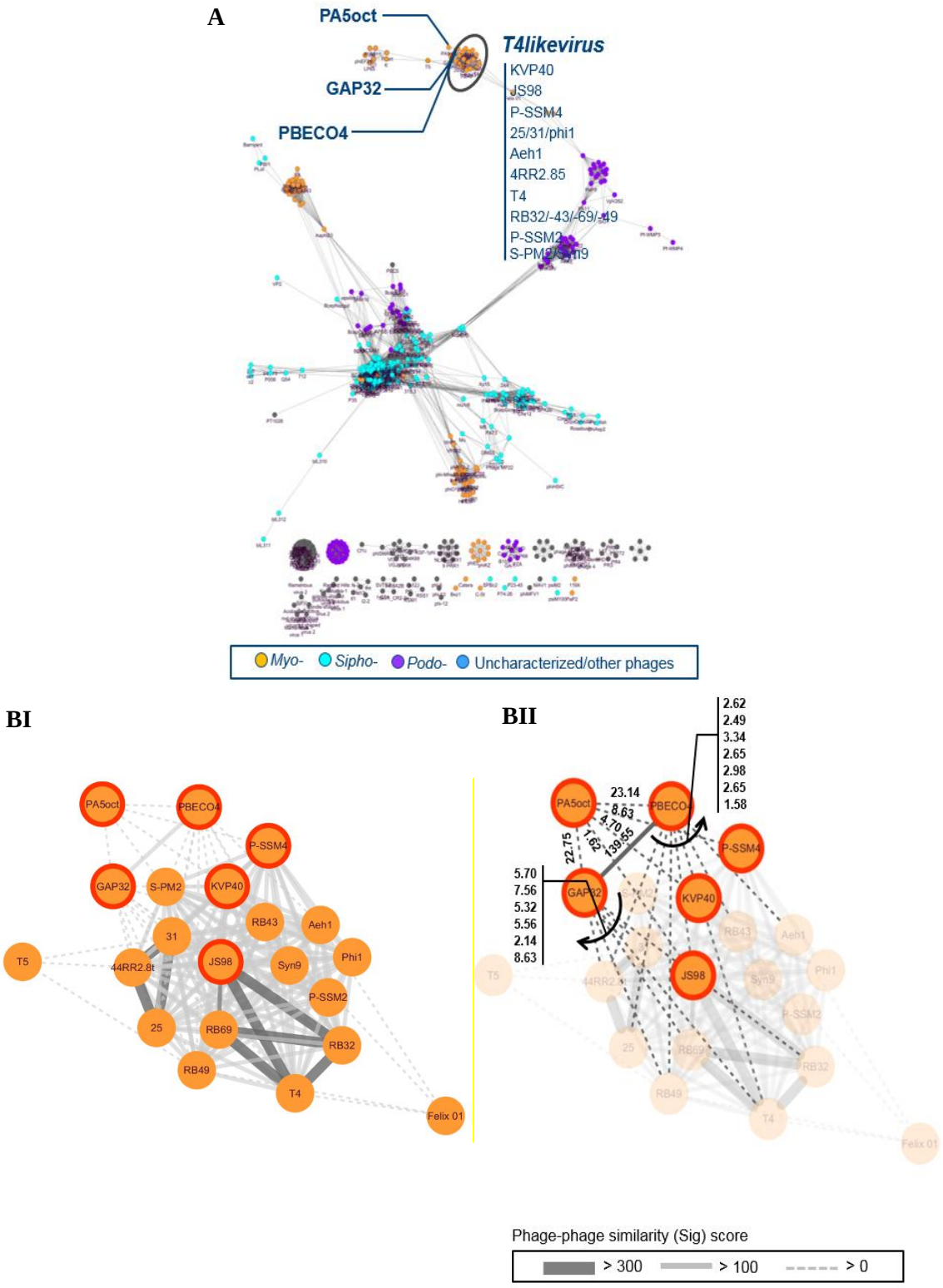


Figure 6.17: Protein-sharing network and reticulate relationship of PA5oct and T4-related phage populations. (A) A network representation was produced using the edge-weighted spring embedded layout of Cytoscape version 3.1.1. Nodes indicate phages and edges between two nodes indicate their statistically weighted pairwise similarities with significance (Sig) scores of >1. Each node is representing bacteriophages belonging to *Myoviridae* (orange), *Podoviridae* (violet), *Siphoviridae* (light blue), or uncharacterized phages (dark blue). Edges between two nodes indicate their statistically weighted pairwise similarities with similarity scores of ≥2. **(B I, II**) Assigning of the phages onto population network composed of interactions among T4-like phages and unclassified *P. aeruginosa* phage PA5oct, *Cronobacter sakazakii* phage GAP32 and *Escherichia* O157-specific phage PBECO4.

Moreover, phages 31 and 25 (*Secunda5virus*), 4RR2.8t (*Biquartavirus*), and phages JS98, RB69, RB32, and T4 (*Tevenvirinae*) showed relatively strong relationships to each other's genomes due to the significant number of shared protein families, indicating large portion of vertically inherited genetic pool (Figure 6.17 A).

PA5oct (novel), PBECO4 and GAP32 (*unclassified Myoviridae*), T5 (*T5virus*), FelixO1 (*Felixo1virus*) appears to be diverged members. Importantly, among these five phages, there are two phages that appear to act as bridge for T4-related components due to their links to other phage groups. These are T5 and FelixO1 (Figure 6.17 B).

General characteristic of PA5oct are presented in table 6.4.

Table 6.4 Major features of PA5oct phage.

Phage	**PA5oct**
Water source	Irrigation field
City/Country/ Date	Wrocław/Poland/ 2011
Host	PAO1
Genome size (bp)	287,182
GC (%)	33.3
ORFs	462
Virion dimensions head/tail [nm] [a]	131/136
Latent period [min]	40
Burst size [pfu/cell]	30-40
Heat stability for 1h [ºC]	4-60
pH stability for 1h at RT	5-11
Chloroform sensitivity at RT/4ºC [h]	24

6.4. Discussion

In this chapter we presented characterization of a novel and unique virus PA5oct, that belongs to the *Myoviridae* family and infects *P. aeruginosa.* It is a member of an emerging group of giant phages, which possibly represents a new genus, based on genome organization and clustering and the existing parameters for taxonomic guidelines (R. Lavigne, personal communication). This group would include coliphage 121Q, *E. coli* 157:H6 phage PBECO4, *Klebsiella* phage RaK2 and *Cronobacter* phage GAP32. The representatives of this group have very large heads, relatively short tails, and limited genomic relationships with other phages (Abbasifar *et al.*, 2014).

In contrast to previously characterized *Pbunaviruses* KT28 and KTN6, as well as φKZ isolate KTN4, PA5oct presented narrow host range and lower stability in a range of temperature and pH. These features could become a potential obstacle in phage preparations for therapeutic purposes. Nevertheless, initial biofilm degradation assessment clearly presented potential of PA5oct for biofilm degradation. This phage was able to create visible tunnels in 24 h and 48 h old biofilm structure. This ability is currently

further evaluated by Drulis-Kawa group and collaborators (Department of Microbiology, Institute of Biology and Department of Molecular Physics, Instituteof Physics, The Jan Kochanowski University in Kielce, Poland). Furthermore, the experiments using the 'Airway Surface Layer' infection model proved its potent antibacterial activity (2.5 - 6.5 logs reduction of an extracellular bacterial load). PA5oct phage was only slightly less efficient compared to the previously characterized KTN4 phage (4 -7 logs reduction of bacterial CFUs) (chapter 5). In the NuLi-1 cell line infection, caused by PAO1 and CF708 strains, KTN4 treatment was more efficient. However, the infection caused by the non-CF0038, burn wound isolate was more susceptible to PA5oct treatment. In case of CuFi-1 infection, treatment with KTN4 reduced greater CFUs of nonCF0038 strain (6 log reduction compared to a 2.5 log reduction by PA5oct), whereas the treatment of PAO1 and CF708 was almost identical for both phages. Morever, substantial differences can be observed during the bacterial internalization process. PA5oct significantly ($p < 0.05$) reduced internalization level of nonCF0038 and CF708 strains into NuLi-1 epithelial cells, by 3 and 1 log respectively, while KTN4 treatment was not effective. The opposite situation was found for the CuFi-1 cell line, where PA5oct treatment was not effective, while KTN4 reduced invasion of nonCF strains by 1 log. These results show, that considering the entire experiment, the efficacy of treatment was highly strain and condition dependent. This analysis confirms that potential phage therapy should be considered in a form of phage cocktails, which would comprise phages with different, complementary features. However, their synergistic effects should be studied in detail. Moreover, both giant phages, KTN4 and PA5oct, could serve as antimicrobial agents, however in altered lung environment their activity could possibly be supported by viscosity reduction agents and/or other antimicrobials.

The PA5oct DNA sequencing revealed a giant linear double-stranded DNA genome, which is the eight largest sequenced myovirus and third largest among *Pseudomonas* phages. Only *P. aeruginosa* phage φPA3 (309 kbp, Monson*et al.*, 2011) and *P. chlororaphis* phage 201φ2-1 (316 kbp, Thomas *et al.*, 2008) possess bigger genomes. Importantly, to date, the largest myovirus known, and also the largest phage in terms of capsid size and DNA content, is *Bacillus megaterium* phage G (head of 160 nm in diameter and 497,513 bp genome). However, some giant phages, especially those reported prior to the 1990s, were identified only by electron microscopy (e.g. *Gluconobacter* phage GW6210) (Abbasifar *et al.*, 2014).

Unfortunately, only for a small number of phage proteins a putative function was predicted. A large pool of ORFs was unique, without any sequence similarity in the currently databases. The highest amino acids similarity was found with *Cronobacter* phage vB_CsaM_GAP32, *Escherichia* phage 121Q and *Enterobacteria* phage vB_KleM_RaK2. Limited sequence similarity may result from low availability of annotated giant genomes. Most databases consist vast pool of genomes < 100 kbp, in which gene features differ greatly compared to giant viruses. For this reason the functional annotation of giant phage genomes is of utmost importance. It will help assess phage safety and explain basic phage biology properties. The biological functions of these pool of unknown proteins could be revealed by e.g. screening the libraries of unknown genes toward e.g. inhibitory, antibiofilm or metabolism modulation functions. However, this research is complex and time consuming (Van den Bossche *et al.*, 2014, Wagemans *et al.*, 2014). Another option is ESI-MS/MS and RNA seq. Mass spectrometry can reveal the presence of proteins involved in virion structure, while RNA seq in general significantly improves genome annotation. Whole genome analysis of transcription using RNA seq is a powerful way to elucidate differential expression of gene features across different conditions. The RNA sequencing has revolutionized genome annotation. By experimentally defining the timing and expression levels of transcripts in both phage and host, directional RNA seq has the ability to discover novel coding sequences, particularly for non-coding RNAs and small phage peptides falling below gene prediction thresholds (Ceyssens *et al.*, 2014). In this manner, we can refine annotations of existing coding sequences, predicted strictly *in silico*, based on the presence of open reading frames and often distant orthology to other often hypothetical features.

To examine the genetic relationships between PA5oct and other jumbo bacteriophages as well as *Myoviridae*, *Siphoviridae*, *Podoviridae* viruses, a protein-sharing network was constructed. The PA5oct, *Escherichia* phages 121Q/PBECO 4, *Klebsiella* phage vB_KleM-RaK2, *Klebsiella* phage K64-1, and *Cronobacter* phage vB_CsaM_GAP32 presented closer relationships based on shared conserved core genes. Furthermore, PA5oct, PBECO4, GAP32, T5, FelixO1 appears to be distantly diverged members of *Tevenvirinae*. The T5 and FelixO1 phages appear also to act as bridge for T4-related components due to their links to other phage groups. Unlike the marker genes of bacteria (i.e.,16S rRNA genes), which can be used for their taxonomic classification, viruses often lack those universal genes. Thus, a traditional single-gene-based phylogeny

can be limited to several viral groups. That is why presented here protein-sharing network introduce another approach, which provides a global view and is consistent with current taxonomic groupings. Using this approach, the possible genetic relationships for not only PA5oct, but all giant phages among whole phage population could be observed. In the network, viral clusters for the giant phages are not always consistent with genus levels as currently defined by ICTV. Specifically, the hypergeometric formula used in this study can measure the significance based on the possibility of gene sharing between genomes, rather than their magnitude, whereas the ICTV uses the magnitude of gene-sharing to define the genus and subfamily (i.e., 40% genus and 20% for subfamily). Thus, some inconsistency can be expected. Presented here connections of PA5oct with other giant viruses as well as their clustering into the same clusters can indicate their closest relationships. Considering all these characteristics, PA5oct phage is an interesting subject to study, which present unique features and should be further explored.

Chapter 7
Characterization of polysaccharide depolymerases domains associated with phage tail apparatus

Contribution

- Experiments were conducted by Katarzyna Danis-Włodarczyk with the help of technician Joris Schuermans and master student Laura Bukenbergs (2016 Master's program at KU Leuven Bioscience Engineering, supervisor Katarzyna Danis-Włodarczyk, promoters: Prof. Rob Lavigne and Dr. Lorena Rodriguez-Rubio).
- The identification and recombinant protein expression of LKA1 gp49 have been presented in doctoral dissertation of Anneleen Cornelissen (Laboratory of Gene Technology, KU Leuven, 2011).
- The LKA1 gp49 crystal structure was provided by Dr. Petr G. Leiman (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland).
- The size exclusion chromatography was performed in the Centre of Microbial and Plant Genetics with the help of technician Jos Desair (Prof. Jan Michiels, KU Leuven).

7.1. Introduction

Highly diverse bacterial polysaccharides are composed of single sugars linked together to complex high molecular weight branching moieties of three or more sugar residues. Based on localization they are classified as intracellular (covalently or ionically associated with the cell peptidoglycan), structural (i.e. LPS) or extracellular (EPS). They play important role in structural cell integrity, cell adherence to surfaces, as well as in biofilm structure (Roach *et al.*, 2015; Campodonico *et al.*, 2008). Exopolysaccharides protect the embedded bacteria from desiccation as well as phagocytosis or antimicrobial agents, promote virulence and mask cells surface receptors, necessary for phage adsorption (Kumar *et al.*, 2007; Looijesteijn *et al.*, 2001; Mushtaq *et al.*, 2005). EPS is the most abundant polysaccharide and can be in form of polysaccharide capsules covalently bound to the cell surface or secreted outside the cell in the form of mucus (Roach *et al.*, 2015).

In nature, specific bacteriophages may possess EPS/LPS depolymerases to degrade bacterial polysaccharides during phage infection. These enzymes can be overproduced and release to the environment during cell lysis or associated to viral tail structure as for example tail spike proteins (Sutherland, 1999; Casjens *et al.*, 2011). The presence of these enzymes is manifested *in vitro* as a halo zone around the phage plaques, which results from the enzymatic degradation of bacterial EPS without phage infection (Azeredo *et al.*, 2008). However, the knowledge about depolymerases is currently very limited and mainly derived from *E. coli* phages (Leiman *et al.*, 2007). Little is known about their substrate binding and cleaving domains. It is also unknown to what extent the depolymerase activity is required for the infection process and how

possible variations in the amount and density of the polysaccharide influence this process (Pelkonen *et al.*, 1992).

The LoGT (KU Leuven) possess a large collection of *Pseudomonas* phages characterized for intended use in the phage therapy of the multidrug resistant bacteria. From this pool two phages with potential virion-associated EPS/LPS degrading enzymes were selected. In this study we isolated the enzyme domains and characterized their structure as well as the expression and purification conditions. The activity of each protein was assessed against bacterial EPS, LPS knock out *Pseudomonas* mutants and biofilms.

7.2. Detection and bioinformatics analysis- a search for polysaccharide depolymerases

For the purpose of this study, two phages, *Luz7virus* LUZ7 and *phiKMVvirus* LKA1 with potential EPS/LPS depolymerases were selected. The presence of these specific enzymes was identified by the formation of opaque-looking halo zones around phage plaques (Fig. 7.1). Diameter of LKA1 halo zone on PAO1 Krylov lawn increased from 0.5 ± 0.1 cm after 24 h, to 0.6 ± 0.2 after 48 h and 0.8 ± 0.1 cm after 72 h incubation at 37°C. In case of LUZ7 the increase in diameter of the halo zone was not observed.

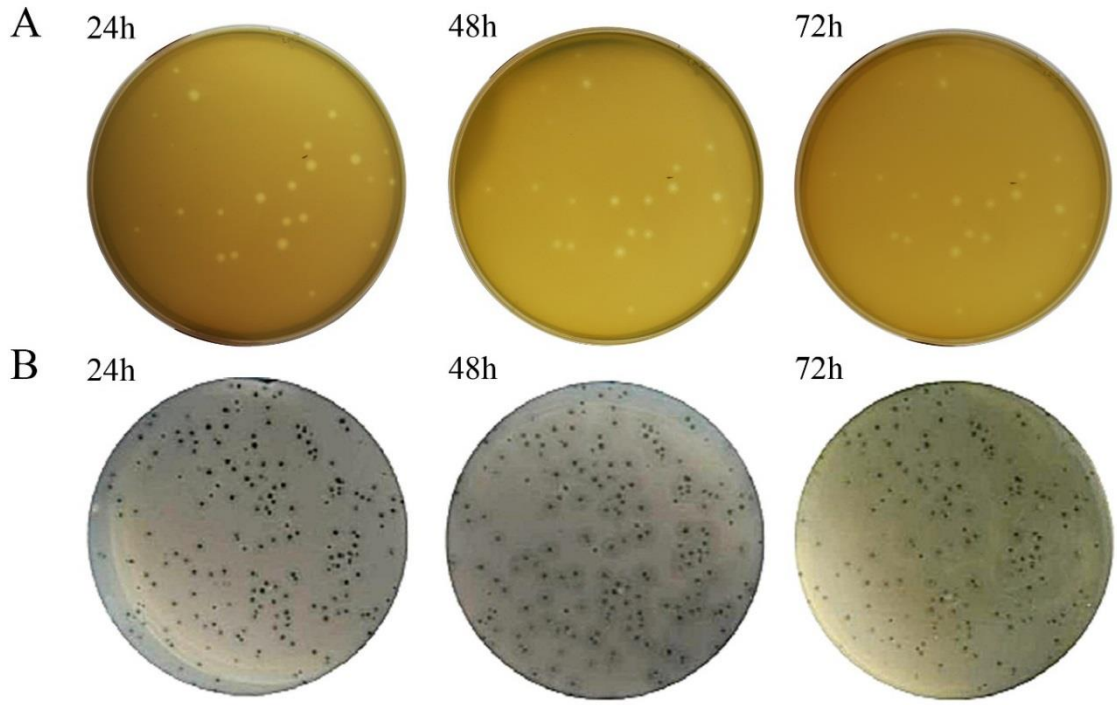


Figure 7.1: Plaque morphology of *Luz7virus* LUZ7 and *phiKMVvirus* LKA1. (A) Plaques of phage LUZ7 and (B) LKA1 upon infection of PAO1 Krylov. In cases of phage LKA1, in the course of time, the diameter of the halo zone surrounding a plaque increases.

After detailed analysis of phage genomes, two large ORFs (LKA1 gp49, 2310 bp; and LUZ7 gp56, 3210 bp) encoding putative polysaccharide depolymerases, were selected for further study. Both genes present only 9.1% similarity to each other (ClustalW2). The LKA1 gp49 was successfully prepared by Anneleen Cornelissen a few years ago (LoGT, KU Leuven, 2011). Unfortunately, the corresponding PCR product of LUZ7 gp56 was very difficult to clone into expression vectors (pEXP5-NT/TOPO®, pEXP5-CT/TOPO®, pBAD/TOPO ThioFusion®) and further express in order to achieve soluble active protein (variation of competent cells and expression conditions, Table 7.3). For this reason, we focused on the depolymerase domains (both LKA1 gp46 and LUZ7 gp56), that in the future could be also suitable candidates for protein engineering (e.g. fusion with other protein domains).

Table 7.1: The list of polysaccharide depolymerases and their domain candidates chosen for this study.

Protein (NCBI code)	Origin	Nucleotide length	Amino acids length	Theoretical pI	Ext. coeff. (M^{-1} cm^{-1})	Putative catalytic specificity	Protein homologue (hhpred, NCBI code, e-value)
LKA1 gp49 (YP_001522890.1)	*P. aeruginosa phiKMVvirus* LKA1	2310	769 (80.34 kDa)	4.77	75470	Pectin/pectate lyase	*Pseudomonas* phage phi297 tail spike gp61 (4RU5, e-value 1.1E-74) (Browning, to be published) *Yersinia enterocolitica* family 28 exopolygalacturonase (2UVF, e-value 1.9E-20) (Abbott *et al.*, 2007)
LKA1 gp49 domain (389 - 584 aa)*		294	198 (20.81 kDa)	4.89	23545		
LUZ7 gp56 (YP_003358338.1)	*P. aeruginosa Luz7virus* LUZ7	3210	1070 (115.74 kDa)	5.32	195985	GDSL-like Lipase/Acylhydrolase family or SGNH hydrolase	*Pseudomonas* phage KPP21 gp96, a putative SGNH hydrolase-type esterase domain protein (e-value 0.0) GDSL-like Lipase from *Parabacteroides distasonis* ATCC 8503 (3P94,e-value 1E-06)
LUZ7 gp56 domain (410 - 673 aa)*		798	266 (29.00 kDa)	6.51	58900		

Abbreviations: (*) amino acid locus in a full length protein; (pI) a isoelectric point; (Ext. coeff.) an extinction coefficient parameter; (NCBI) National Center for Biotechnology Information.

.

LKA1 gp49

Gp49 is a tail spike protein of *phiKMVvirus* LKA1. Bioinformatics analysis revealed that its C-terminal part (136 - 769 aa) consist of a right-handed β-helical structure elements typical for carbohydrate-binding proteins, recognizing LPS or EPS structures (Fig. 7.2), and often serves as the determinant of phage host cell specificity (Molineux, 2006; Barbirz *et al*., 2008; Muller *et al*., 2008; Walter *et al*., 2008). It is similar to the C-terminal region of *P. aeruginosa* siphovirus D3 gp27 (Kropinski, 2000). The external part of the C-terminus has no amino acid sequence or secondary structure homology. In other tail spike proteins, this C-terminal part have been shown to be important in the folding and trimerization processes (Schwarzer *et al*., 2007; Walter *et al*., 2008; Xiang *et al*., 2009). A centrally located putative pectate lyase domain (Phyre1 136 - 769 aa, Phyre2 180 - 681 aa, InterPro 222 - 546 aa, HMMER 389 - 584 aa, HHpred 223 - 559 aa) was also identified. The pectate lyase superfamily is important in galacturonan components degradation of pectin (Jenkins *et al.*, 2001). The conserved N-terminus of LKA1 gp49 is similar to N-terminal region of the T7 phage tail spike protein gp17 and provides the stable association of the catalytic C-terminal region with the tail structure (Steven *et al.*, 1988).

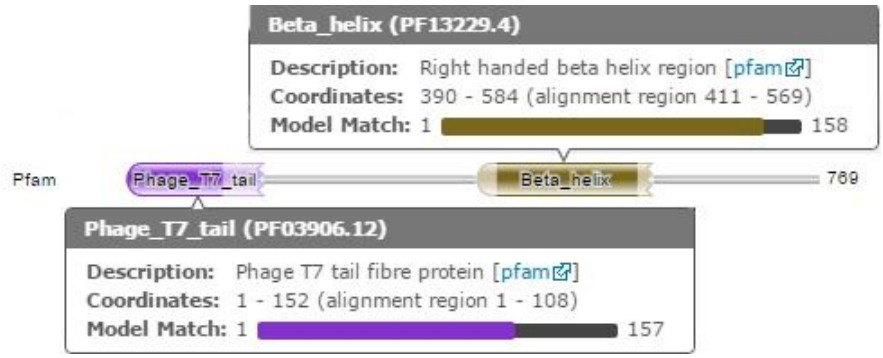


Figure 7.2: Pfam analysis of LKA1 gp49 structure. Predicted conserved N-terminus associated with tail structure is marked violet, C-terminal domain with catalytic activity marked olive (Hmmer).

The crystal structure of LKA1 gp49, lacking the highly flexible N-terminal phage tail fiber region was resolved in a collaboration with Dr. Petr G. Leiman (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland) (Fig. 7.3). The crystallography experiments elucidated a hexamer composed of two homo trimer A3 structures, built of six chains, A-F, (resolution of 1.9 Å, R-factor 0.132, R-free 0.174, PDB code 4RU4).

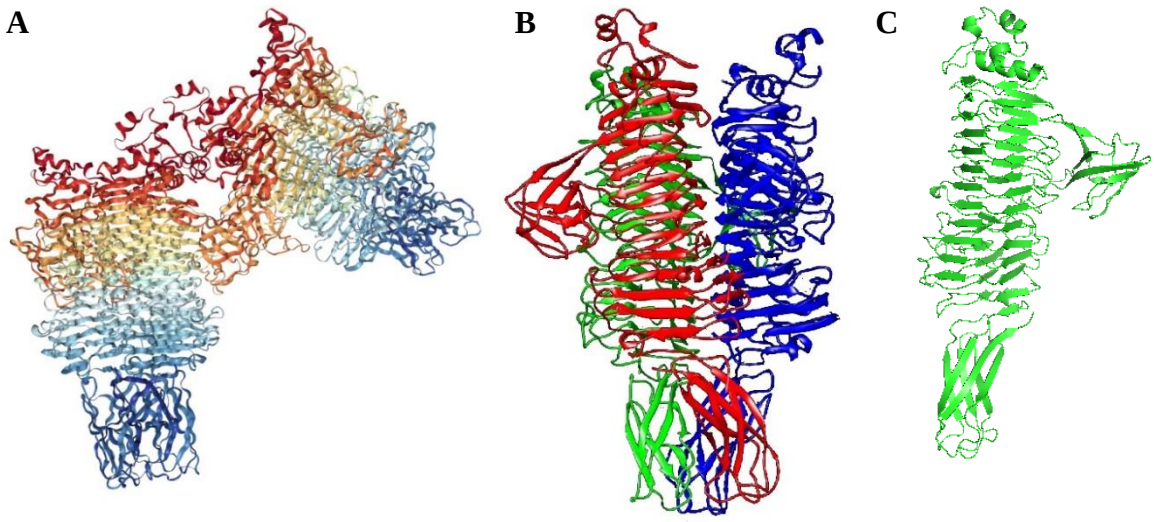


Figure 7.3: Representation of three dimensional structure of LKA1 gp49 based on crystallography. (A) hexamer composed of two trimeric structures. (B) trimeric structure composed of three chains, marked red, blue, green. (C) single chain. Cartoon representation, PyMol modelling.

Further structure modelling in PyMol and MOE programs allowed us to identify the localization of enzyme active center (Fig. 7.4). This area constitutes a group of amino acids which are close to each other in a three-dimensional molecule. They are responsible for the binding of substrates and the prosthetic groups to form a complex of enzyme-substrate, wherein the reactive moieties are located near the center of the catalyst. This may occur with changes in the conformation of both, the substrate and enzyme.

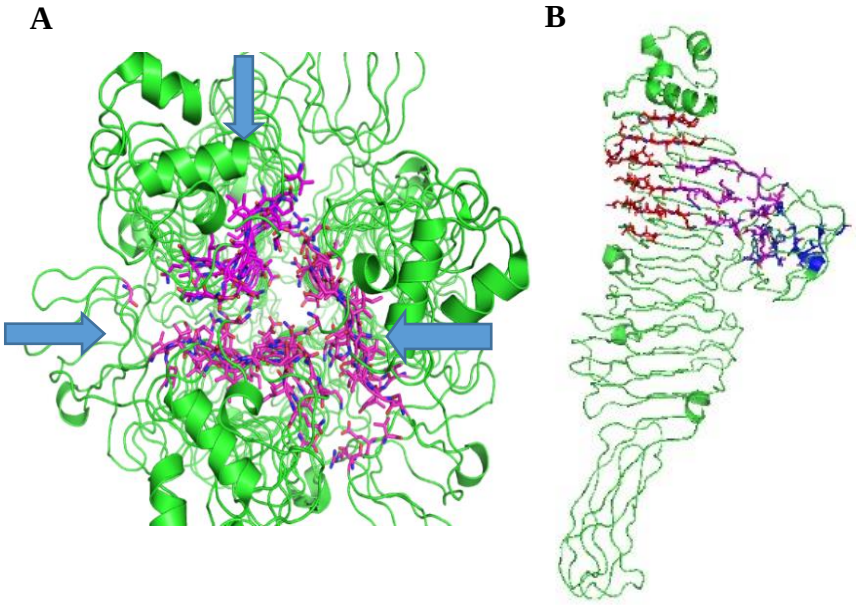


Figure 7.4: Crystal structure of LKA1 gp49. (A) schematic representation of LKA1 gp49 active center in protein trimer. Blue arrows and purple sticks represent amino acid sequences belonging to enzyme active site in the individual trimer subunits, marked as green cartoons. View from top of trimer. (B) schematic representation of LKA1 gp49 active center in the single chain. Red, purple and blue sticks represent three sites of active center. Green cartoons represent structure of single subunit. Side view. PyMol modelling.

MOE analysis allowed for prediction of three sites of enzyme active center, that could in the future be targets of a side-directed mutagenesis (Fig. 7.5).

Site 1 (168-454 aa)

GSVGQSLQFLEMGRVTPAQFGAVGDGASHPLSERYATLAEAQTVYPHAVALSDEIDWAALQAAVDSGA**PVH**IPSGDYQINRG
ISSTGSL**QIAGDGA**TSIIRPTAAFTGTSVLSCVGSLVALPNISSVSAGSLTIDFASTPNLVA**G**DVFIIYNPTDSSFSGFRTSYRAGEFCEV
RAVSGNTVTIRSALYAAYDGATVAIYKVV**SGVVD**I**ASIQ**IVGGTVPMNGLLVEA**VVSPRVDDVT**VTLANNAGVYFARCYDA**KIT**
NSNISNIGDGGDDYGIIFGNCHDGGA**DN**C**K**

Site 2 (289-470 aa)

P**NIS**SVSAGSLTIDF**A**STPNLVAGDVFII**YNPTD**SSFSGFRTSY**R**AGEFCEVRAVSGNTVTIRSALYAAYD**GATVAIY**KVVSGVVDIA
SIQIVGGTVPMNGLLVEAVVSPRVDDVTVTLANNAGVYFARCYDAKITNSNISNIGDGGDDYGIIFGNCHDGGADNCKVYARRHAI
ATGGD**A**E**V**

Site 3 (193-454 aa)

GASHPLSERYATLAEAQTVYPHAVALSDEIDWAALQAAVDSGAPVHIPSGDYQINRGISSTGSLQIAGDGATSIIRPTAAFTGTSVLS
CVGSLVALPNISSVSAGSLTIDFASTPNLVA**GDV**FIIYNPTDSSFSGFRTSYRAGEFC**E**VRAVSGNTVTI**R**SALYAAYDGATVAIY**KV**
VSGVVDIA**SIQ**IVGGTVPMNGLLVEA**VVSPR**VD**DVT**VTLANNAGVYFARC**YD**AKIT**NSNIS**NIGDGGDDYGIIFGNCHDGGADNC
K

Figure 7.5: Three sites of enzymatic active center of LKA1 gp49 predicted with MOE. Bolded amino acids are predicted to play the key role in the active center.

The secondary structure analysis revealed, that a single chain is composed of: 6% helical (8 helices; 37 residues) and 48% β-sheet (15 β-bridges and 47 β-strands; 291 residues) structures. The results are summarized on figure 7.6. This analysis was crucial to properly defined LKA1 gp49 domain, without cutting secondary structures, i.e. helices and strands.

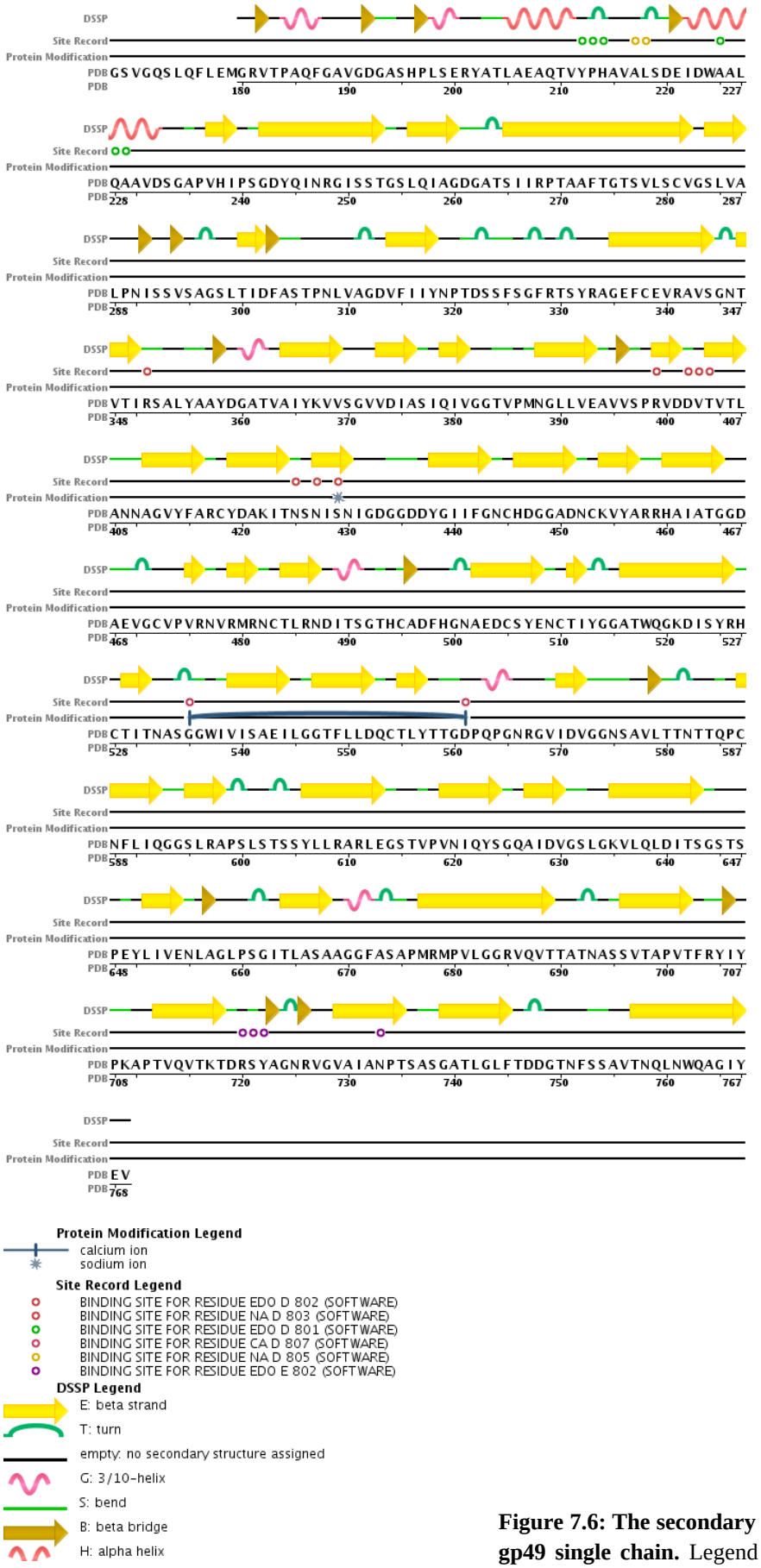


Figure 7.6: The secondary structure analysis of LKA1 gp49 single chain. Legend is provided below diagram. PDBsum analysis.

Based on above analysis LKA1 gp49 domain was chosen for further analysis (Fig. 7.7, Table 7.1). Additional margin of amino acids was added to prevent cuts in the secondary structures, i.e. helices and strands, that could have an impact on protein folding.

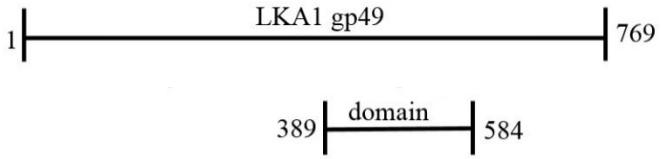


Figure 7.7: Prediction of LKA1 gp49 domain. From top to bottom: whole length protein and domain (389 - 584 aa, pfam). Amino acid position is given on the left and right side.

LUZ7 gp56

Gp56 is a tail fiber protein of *Luz7virus* LUZ7 with centrally located putative GDSL-like Lipase/Acylhydrolase family (SGNH hydrolase) domain (453 - 653 aa, BLAST P, hhpred, InterPro Scan; 459-641 aa pfam) (Fig. 7.8).

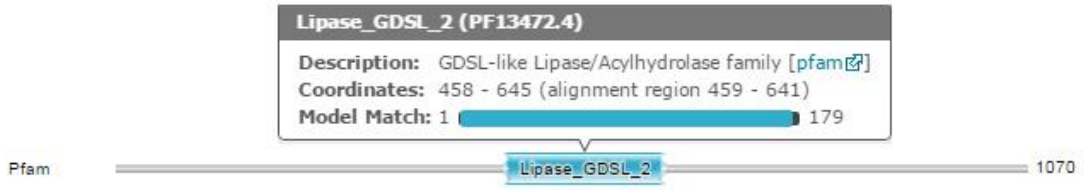


Figure 7.8: Pfam analysis of LUZ7 gp56 structure. The GDSL-like Lipase/Acylhydrolase family (SGNH hydrolase) domain is marked blue (Hmmer).

GDSL lipases are hydrolytic enzymes with multifunctional properties, such as broad substrate specificity and regiospecificity. All representatives have five conserved sequence blocks (I–V) and are further classified based on the strict conservation of the catalytic residues Ser, Gly, Asn, and His within conserved blocks I, II, III, and V, respectively (Fig. 7.9) (Dalrymple *et al.*, 1997; Ling *et al.*, 2006; Mølgaard *et al.*, 2000). They exhibit a GxSxxxxG motif, in which the active site serine is located near the N-terminus and display a Gly-Asp-Ser-(Leu) [GDS(L)] motif in conserved block I (Jiang *et al.*, 2012; Brick *et al.*, 1995; Lee *et al.*, 2009; Updegraff *et al.*, 2009; Upton *et al.*, 1995). These enzymes have little sequence homology with true lipases and have a similar fold to flavoproteins, i.e. a three-layer α/β/α structure, where the β-sheets are composed of five parallel strands (Akoh *et al.*, 2004).

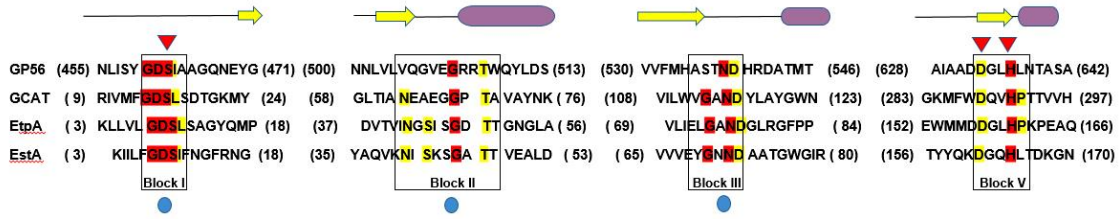


Figure 7.9: Sequence alignment of four conserved blocks in the GDSL-like Lipase family. Sequence alignment of conserved blocks and flanking residues in LUZ7 gp56 (GP56); GCAT: *Aeromonas hydrophila* lipase/acyltransferase (SWISS-PROT P10480); EtpA: *Vibrio mimicus* arylesterase (SWISS-PROT Q07792); EstA: *Lactobacillus helveticus* CNRZ32 arylesterase (TrEMBL Q9LAH7). Four consensus blocks I, II, III, and V are boxed by a black line. Conserved residues are masked in red (absolutely conserved) or yellow (in at least two sequences). The number in parentheses indicates the location of the end residue. The secondary element assignment, above the alignment, corresponds to the LUZ7 gp56 structure. Yellow arrow corresponds to strand, violet cylinder to helix. The catalytic triad and oxyanion hole residues are indicated by red-filled down triangles and blue-filled circles, respectively.

Analysis of secondary structure revealed, that LUZ7 gp56 domain is composed of 8 α-helices, 4 β-strand, 48 β-turns and 6 γ-turns. The results are summarized in figure 7.10. As previously described, this analysis was crucial for precise definition of domain length.

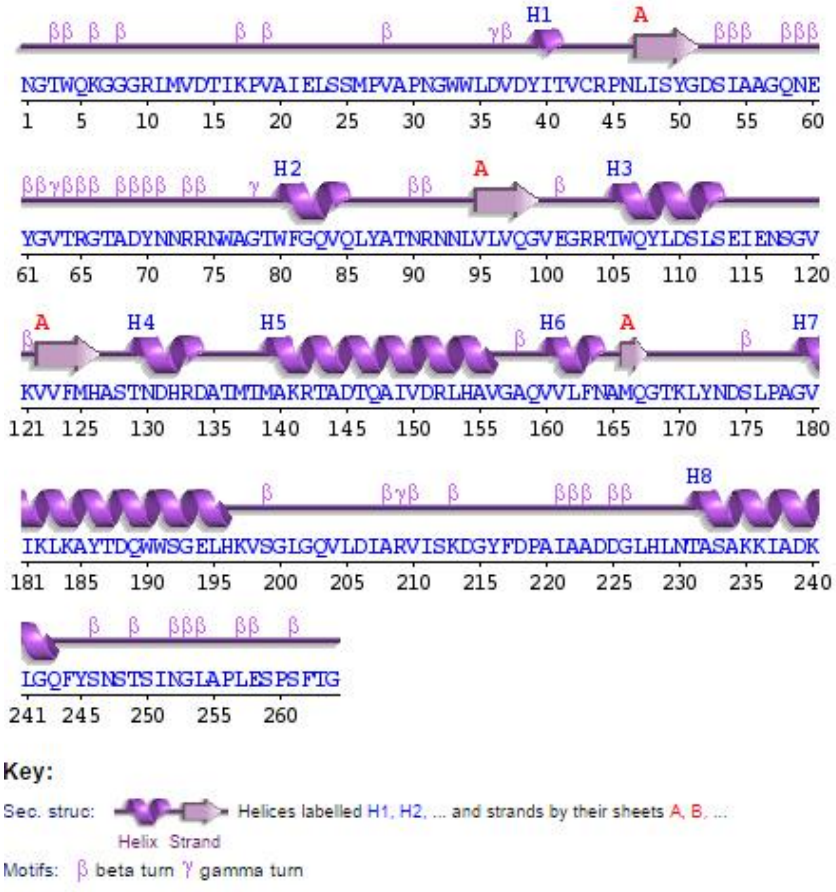


Figure. 7.10: The secondary structure analysis of LUZ7 gp56. Legend is provided below diagram. PDBsum analysis.

The tertiary structure predictions of the whole length LUZ7 gp56 showed low homology (Phyre2). However, when we focused only on the centrally located domain with additional margin of amino acids estimated based on secondary structure prediction (410 - 673 aa), the tertiary structure could be predicted with the highest homology corresponding to Flavodoxin-like SGNH acetylhydrolase (98.2% confidence in Phyre2) (Fig. 7.11).

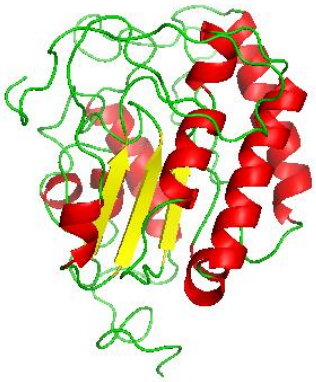


Figure 7.11: Prediction of LUZ7 gp56 domain tertiary structure. Tertiary structure with loops and turns marked green, β-sheets yellow and α-helices red. Cartoon representation. PyMol modelling.

The ligand binding site could be predicted with the use of 3DLigandSite-Ligand binding site prediction Server (Table 7.2, Fig. 7.12). The GDSL lipases have a flexible active site, which alters its conformation in the presence and binding of the different substrates (Akoh *et al.*, 2004; Ling *et al.*, 2006).

Table 7.2: Prediction of the LUZ7 gp56 domain ligand binding site based on structure homology (3DLigandSite-Ligand binding site prediction Server).

Protein	The closest protein structure homologue	Amino acid composition of ligand binding side
LUZ7 gp56 domain	crystal structure of bovin brain platelet activating factor acetylhydrolase covalently inhibited by soman (85% coverage, PDB code 3DT9, Epstein *et al.*, 2009)	52 ASP, 53 SER, 54 ILE, 60 GLU, 100 VAL, 101 GLU, 102 GLY, 103 ARG, 104 ARG, 126 HIS, 129 THR, 130 ASN, 134 ASP, 170 LYS, 171 LEU, 172 TYR, 173 ASN, 210 VAL, 217 PHE, 218 ASP, 219 PRO, 224 ASP, 225 ASP, 226 GLY, 227 LEU, 228 HIS, 229 LEU, 233 SER

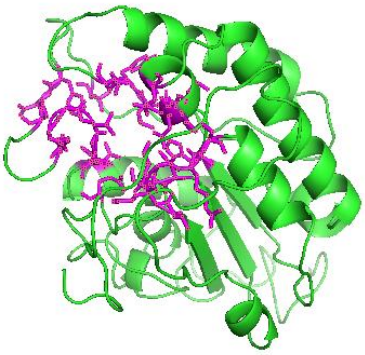


Figure 7.12: Predicted binding site of the LUZ7 gp56 domain. The binding site marked with purple sticks, the rest of the protein presented as a green cartoons. PyMol modelling.

For further analysis one domain was selected (410 - 673 aa), based on analysis presented above (Fig. 7.13).

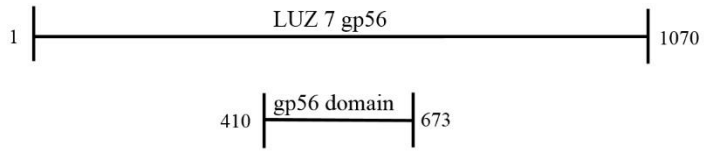


Figure 7.13: Representation of selected LUZ7 gp56 domain. Amino acid position is given on the left and right side.

7.3. Cloning, recombinant expression and purification

The gene sequences encoding two putative EPS/LPS depolymerases and their domains (Table 7.1) were amplified by Phusion® High-Fidelity DNA polymerase with low error rate. First, all PCR products were cloned into pEXP5-CT/TOPO® or pBAD/TOPO ThioFusion® vector, however protein expression did not occurred or proteins were poorly expressed in soluble fraction with the presence of inclusion bodies (0.06 - 0.21 mg/ml). Changing conditions, such as competent cells [*E. coli* BL21(DE3)pLysS, BL21 A1, *E. coli* Lemo21(DE3), *E. coli* BL21-Codon Plus(DE3), *E. coli* ArcticExpress(DE3), Rosetta(DE3)], medium (LB, modified LB, TB, 2xTY or autoinduction), expression temperature (37°C or 16°C) or inducer concentration (1M - 0.1 mM) did not improved protein expression. Finally, the expression vector was changed to pEXP5-NT/TOPO® with a 6x His-tag in the N-terminus and expression was successful, except full length LUZ7 gp56 that still remained unexpressed. The optimal conditions, i. e. where protein is the most abundant in the soluble fraction, were as follows: 1) *E. coli* BL21(DE3)pLysS, 2xTY medium, 0.1 mM IPTG and overnight incubation at 16°C for LUZ7 gp56 domain; 2) *E. coli* Lemo21(DE3), 2xTY medium, 0.4 mM IPTG and 0.5 mM L-rhamnose for full length LKA1 gp49 (prepared from A. Cornelissen construct) and its domain. A summary of the expression optimization is presented in table 7.3.

Table 7.3: Protein expression optimization process for four putative EPS/LPS degrading enzymes. Gray color marks conditions were expression was successful and protein appeared in soluble fraction, however not in all cases expression yield was sufficient.

Type	Name	LKA1 gp49		LUZ7 gp56	
		gp49	domain	gp56	domain
vector	pEXP5-CT/TOPO			√	√
	pEXP5-NT/TOPO	√	√	√	√
	pBAD/TOPO ThioFusion	√	√		
expression strain	Lemo21 (DE3)	√	√	√	√
	BL21 CodonPlus (DE3)			√	√
	BL21 (DE3) pLysS	√	√	√	√
	BL21 AI	√	√	√	√
	Rosetta(DE3)				√
	Arctic Express (DE3)				√
medium	LB	√	√	√	√
	2xTY	√	√	√	√
	Modified LB	√	√		√
	TB	√	√		√
	Autoinduction	√	√		√
temperature post induction	37°C (4.5 h)			√	
	16°C (overnight)	√	√	√	√
L-rhamnose	0.5 mM	√	√	√	√
IPTG	0.1 mM	√	√	√	√
	0.4 mM	√	√	√	√
final conditions		pEXP5-NT/TOPO Lemo21 (DE3) LB medium 0.4 mM IPTG 0.5 mM L-rhamnose 16°C (overnight expression)		----	pEXP5-NT/TOPO *E. coli* BL21 (DE3)pLysS, 2xTY medium, 0.1 mM IPTG 16°C (overnight expression)

The presence of recombinant proteins was checked after expression with the use of SDS-PAGE and Western Blot methods (Fig. S 7.1 - 7.3). It also has to be stressed, that expression of LUZ7 gp56 domain was still difficult and inconsistent with presence of inclusion bodies. Moreover, despite keeping all the time the same expression conditions, protein was not always active, probably due to an incorrect fold or a poor stability after purification.

Finally, a large scale expression (500 ml culture) was performed and proteins were purified with the use of ÄKTA FPLC system in combination with HisTrap^{TM} HP 1 ml columns (Fig. S 7.1 - 7.3). The purity of all proteins was further increased by filtration through 10 K or 30 K filter units, depending on the protein size, and the size exclusion chromatography (Fig. S 7.1 - 7.3). Pure proteins were kept in elution buffer at 4°C with precipitation inspections.

Based on the molecular weight of the fraction eluted after exclusion chromatography, the theoretical size of pure protein could be determined in native state. This could be accomplished with the use of column calibration curve of the HiLoad 16/600 Superdex 200 Prep grade gel filtration column (GE Healthcare) and Superdex 200HP GL10/30 gel filtration column (GE Healthcare). In the case of LKA1 gp49 there was a protein peak present at 59.31 ml of elution, which corresponds to 286.70 kDa. Unfortunately, this protein peak was out of range of column and cannot be precisely assessed. However, this result suggest that protein (80.34 kDa) creates a trimer. This was also observed for other polysaccharide-degrading tail spike proteins interacting with LPS-structures (Miller *et al*., 1998; Freiberg *et al*., 2003; Walter *et al*., 2008) or EPS-molecules (Muhlenhof *et al*., 2003; Stummeyer *et al*., 2006). The crystal structure of LKA1 gp49 confirms this result. The chromatogram of LKA1 gp49 domain was highly unusual. Seven protein peak were registered at 46.25, 57.59, 69.68, 83.66, 86.65, 99.16 and 110.22 ml of elution, which corresponds to 872.27, 331.95, 118.51, 36.01, 27.92, 9.62 and 3.75 kDa, respectively. First two protein peaks were also out of range of column and cannot be precisely assessed. Protein (20.81 kDa) is the most abundant in the second protein peak, and later in first, third and fourth protein peak, the rest of the protein peaks represented expression impurities. This means that LKA1 gp49 domain possibly aggregates. Further analysis is required using e.g. Multiangle Light Scattering – SEC or Circular Dichroism methods to determine if the domain is properly folded and what kind of multimer is formed. Finally, the chromatogram of LUZ7 gp56 domain recorded 4 protein peaks at 11.65, 14.34, 15.66 and 16.55 ml of elution, which corresponds to 454.20; 123.16; 64.91 and 42.15 kDa. First protein peak is out of range of column and cannot be precisely assessed, however gp56 domain (29.00 kDa) was the most abundant in that protein peak. The 12% SDS-PAGE gel analysis of elution fractions of that protein peak elucidated a band on the weight corresponding to dimer of this domain or the protein could aggregate (Fig. S 7.3). Further analysis would be also required. The protein expression yields are presented in table 7.4. The highest concentration was achieved for LKA1 gp49 (1.36 g/500ml).

Table 7.4: The average yields of purified proteins. Measurements were performed on a Nanodrop at wavelength of 280nm with specific parameters of molecular weight and extinction coefficient (Expasy Protparam).

Protein	Total protein yield (g/ 500 ml) after FPLC affinity purification	Total protein yield (g/ 500ml) after size exclusion purification
LKA1 gp49	7.64	1.36
LKA1 gp49 domain	0.97	0.1
LUZ7 gp56 domain	0.99	0.2

(FPLC) Fast Protein Liquid Chromatography

7.4. Activity assessment of polysaccharide degrading enzymes

To assess recombinant protein activity first an elution buffer with imidazole was exchanges to PBS pH 7.5 with the use of 10 K or 30 K Amicon filter units. Next, 10 µl of pure recombinant protein from the highest purification elution fraction was dropped on a *P. aeruginosa* PAO1 Krylov lawn (Fig. 7.14).

The recombinant LKA1 gp49 and its domain preparations (Fig. 7.14 A) created an opaque-looking halo zones (> 2.0 cm), which diameter increased upon incubation at 37°C. For LUZ7 gp56 no enlargement in diameter was observed and area of spotting was unusual, resembling rather phage plaques with resistant colonies on it (Fig. 7.14 B). Protein was manufactured two more times, but the same effect was achieved or there was no activity. It is required to repeat this protein preparation and activity evaluation.

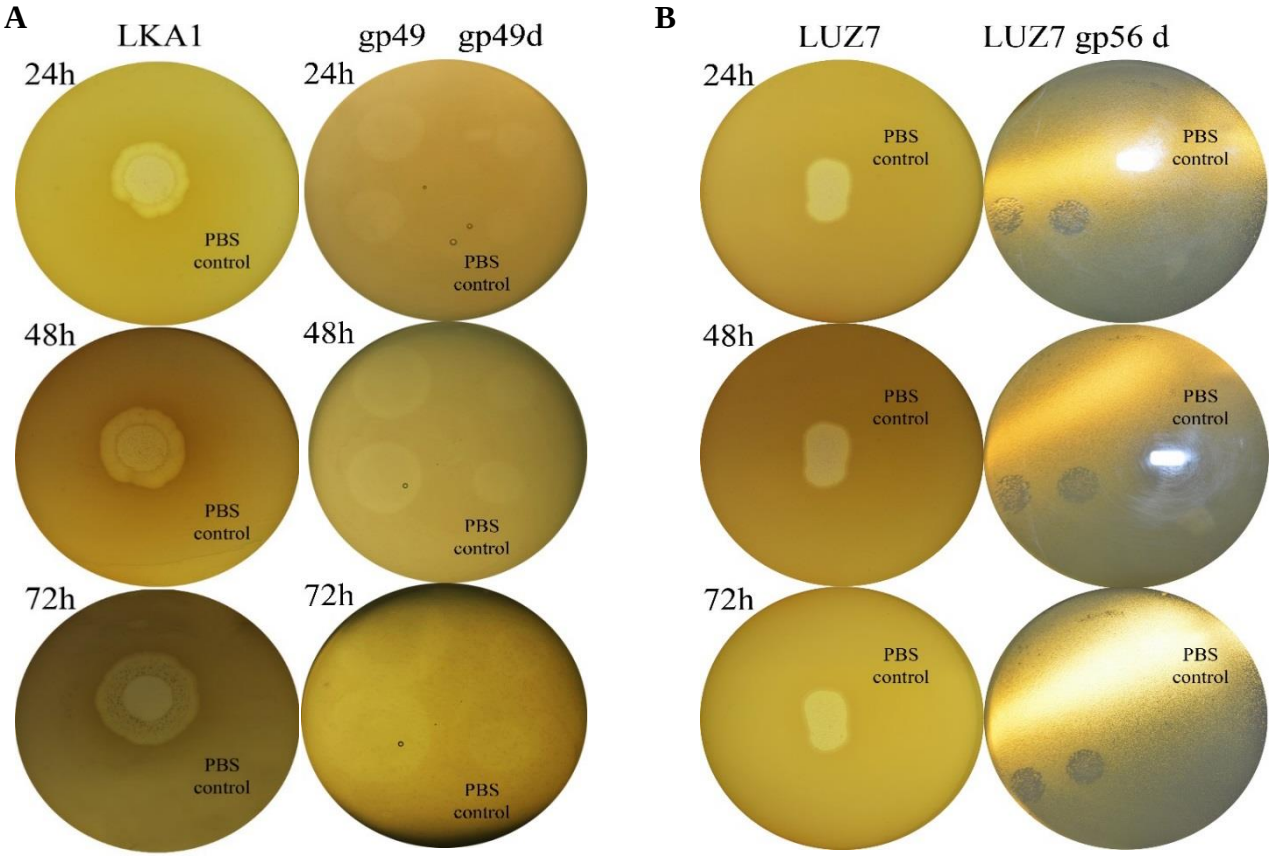


Figure 7.14: The activity assessment of potential polysaccharide depolymerases on *P. aeruginosa* PAO1 Krylov lawn. In all plates 10 µl of sample was spotted. **(A)** From left to right: 1) CsCl pure LKA1 phage solution (8.84 x 10^8 pfu/ ml) with PBS control below visible spot; 2) pure recombinant protein of LKA1 gp49 (2.00 mg/ml, 24.89 µM, left halo zone) and LKA1 gp49 domain (0.5 mg/ml, 24.03 µM) with PBS control below visible spots; **(B)** From left to right: 1) CsCl pure LUZ7 phage solution (1.39 x 10^9 pfu/ ml) with PBS control below visible spot; 2) 1 mg/ml (34.48 µM) LUZ7 gp56 domain (2 repeats) with PBS control above visible spot. The size of each phage plaque and spot area is comparable. In case of phage LKA1 gp49 and its domain diameter of halo zones increases upon incubation at 37°C.

LKA1 phage uses the LPS B-band O-antigen as a primary host cell receptor. LPS is the outermost polysaccharide molecule of *P. aeruginosa* PAO1 outer membrane. Based on this information, putative pure recombinant proteins were spotted (10µl, same protein concentrations as described above) on *P. aeruginosa* PAO1 LPS mutants lawn (Table 7.5). Proteins created halo zone only in case of PAO1 Δrmd (A-, B+) mutant, which suggest that they are able to degrade B-band of PAO1 LPS.

Table.7.5: Recombinant EPS depolymerases halo formation on *P. aeruginosa* PAO1 LPS mutants.

Strain	Feature	LKA1		LUZ7 gp56 domain
		gp49	domain	
PAO1 Δrmd (A-, B+)	Deficiency in D-rhamnose biosynthesis; lack of A-band LPS	H	H	H
PAO1 ΔrmLC (A-, B-, core-)	Deficiency in L-rhamnose biosynthesis; truncate core region, lack of A-band and B-band LPS	/	/	/
PAO1 ΔwaaL (A-, B-)	Lack of WaaL ligating O-polymer to core-lipid A; LPS is devoid of A-band and B-band, semirough (SR-LPS, or core-plus-one O-antigen)	/	/	/
PAO1 ΔwbpL (A-, B-)	Lack of glucosyltransferase WbpL essential for initiation of both A-band and B-band synthesis	/	/	/

(H=halo formation).

Further microscopic analysis of halo zones and the outside bacterial zone was performed (Fig. 7.15). It was clearly showed that EPS material keeps cells closely associated with small clusters. Bacteria taken from halo zone were separated from each other and their EPS slime was reduced.

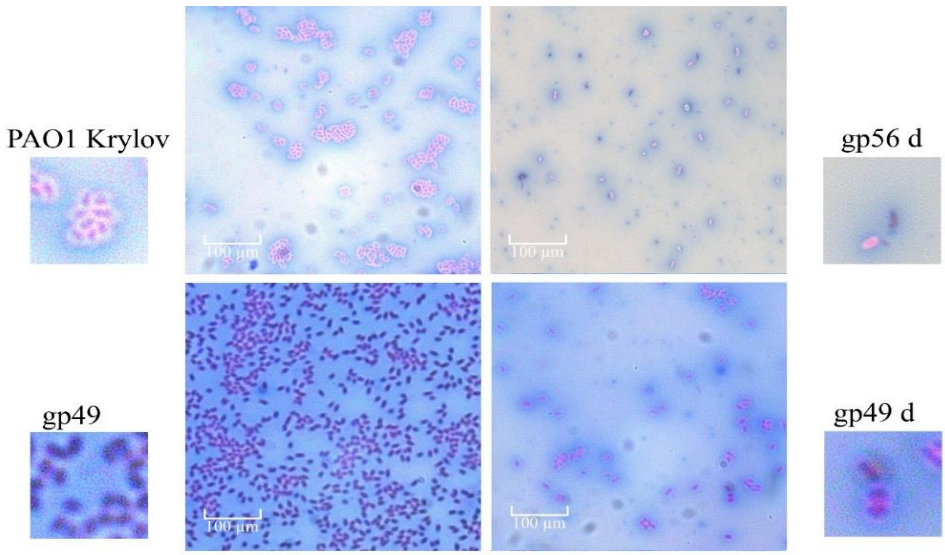


Figure 7.15: Negative staining of *P. aeruginosa* PAO1 cells in the outside bacterial and in the halo zone. Within the outside bacterial zone (top left) bacteria are surrounded by withe EPS slime and backed closely together, while in the halo zone cells are spread separately and the surrounding EPS was reduced. Scale bar represents 100 µm. Under protein name are presented enlarged bacterial cells images.

Finally, to assess the biofilm-degradative properties of EPS/LPS depolymerase domains, a 24 and 48 h old biofilms of *P. aeruginosa* PAO1 Krylov were formed on peg lid plates under static conditions (Fig. 7.16). The incubation temperature was 37°C and the treatment time with protein domains was 30 min. Both conditions were demonstrated to be most suitable for the activity in case of other depolymerases reported so far, e.g. Dpo7 (Gutiérrez *et al.*, 2015). All the assays were performed twice, with four repeats for each. As a negative control served biofilm treated with LB broth [(-) LB]. The positive control was a treatment with high concentration (837.53 μM, 0.4 mg/ml) of gentamicin. The gentamicin MIC for the tested strain was 0.5 μg/ml (1.05 μM). The final protein domains concentrations, chosen based on the expression efficiency, were as follows: 1) LKA1 gp46: 2 mg/ml (24.89 μM); 1 mg/ml (12.44 μM), 0.5 mg/ml (6.22 μM); 2) LKA1 gp46 domain: 0.1 mg/ml (4.80 μM), 0.05 mg/ml (2.40 μM) and 0.025 mg/ml (1.20 μM); 3) LUZ7 gp56 domain: 0.07 mg/ml (2.4 μM) and 0.03 mg/ml (1.2 μM) (Fig. 7.16). For statistic assessment a standard t-tests ($p < 0.05$) were performed.

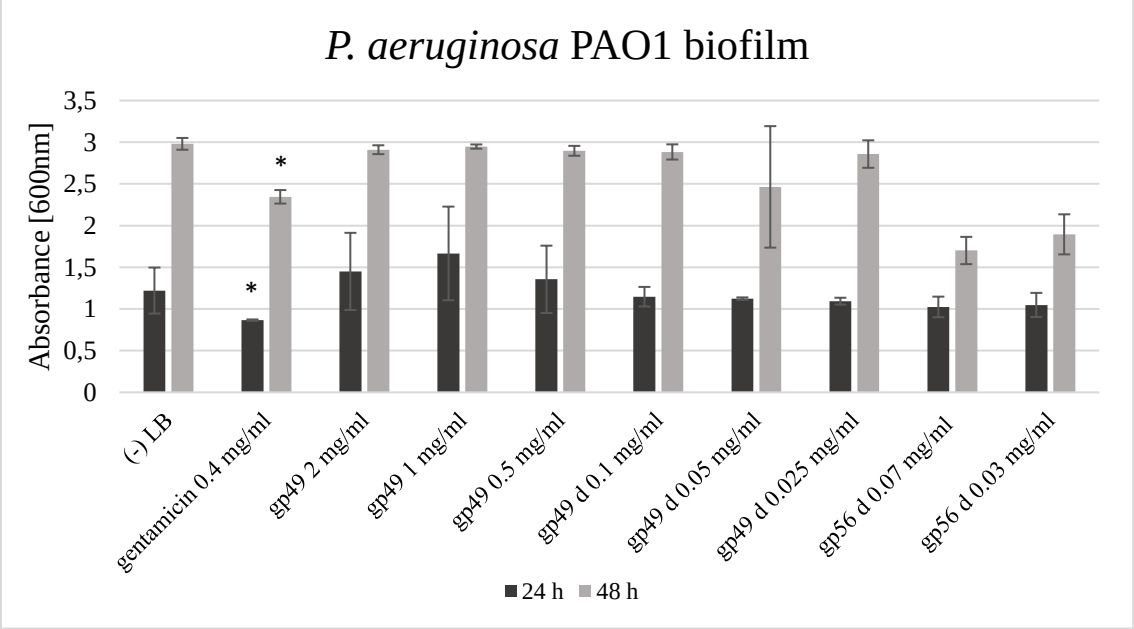


Figure 7.16: The antibiofilm effect of LKA1 gp49 and its domain and LUZ7 gp56 domain 24 h and 48 h *P. aeruginosa* PAO1 biofilm formed on peg-lid plates. As a negative control serves biofilm treated with LB broth [(-) LB], as a positive gentamicin treatment [(+) gentamicin (0.4 mg/ml, 837.53μM)]. Further are presented activities of domains and their two-fold dilutions: 1) LKA1 gp46: 2 mg/ml (24.89 μM); 1 mg/ml (12.44 μM) and 0.5 mg/ml (6.22 μM); 2) LKA1 gp46 domain: 0.1 mg/ml (4.80 μM) and 0.05 mg/ml (2.40 μM) and 0.025 mg/ml (1.20 μM); 3) LUZ7 gp56 domain: 0.07 mg/ml (2.4 μM) and 0.03 mg/ml (1.2 μM). The results are presented as the means ±SD. (*) statistical significant biofilm reduction.

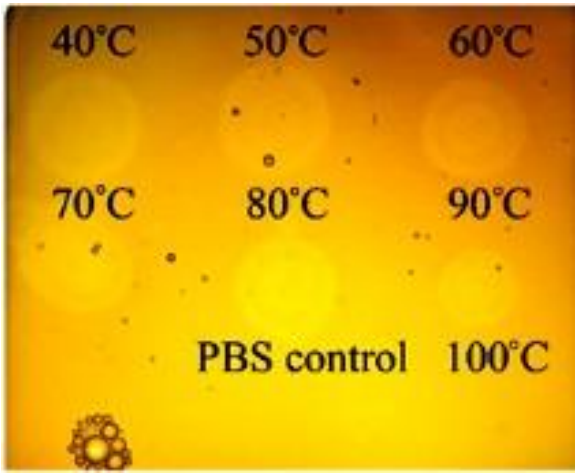
In the case of 24 h old biofilm treatment there was a significant ($p < 0.05$) difference between negative control and gentamicin. The high dose of antibiotic treatment was able to reduce biofilm by 29%. However, the protein treatments had no influence

on biofilm formation. The gentamicin treatment presented also a significant ($p < 0.05$) degradation of 48 h old biofilm, (21% reduction), whereas all the rest recombinant proteins again did not reduced biofilm matrix.

7.5. Protein stability

Since LKA1 gp49 domain presented the strongest activity against *P. aeruginosa* PAO1 polysaccharides, its stability was further evaluated in compare to full length gp49. After 1h incubation at temperature range of 40 - 100 °C, LKA1 gp49 preserved its activity at temperature range of 40 - 90°C (Fig. 7.17 A). In contrast, LKA1 gp49 domain was less stable and preserved its activity at temperature range of 40 - 70 °C. After 1 h incubation at 70 °C, the halo zone is clearly weaker and less visible (Fig. 7.17 B).

A



B

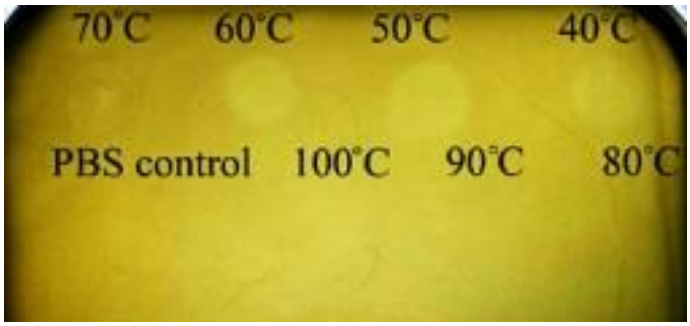


Figure 7.17: Thermal stability of LKA1 gp49 and its domain. A 10µl smple of (A) LKA1 gp49 (0.5 mg/ml, 6.22 µM) and (B) LKA1 gp49 domain (0.1 mg/ml, 4.80 µM) was spotted on *P. aeruginosa* PAO1 Krylov lawn. Samples were spotted after 1h incubation at temperature range of 40 - 100°C.

Further, the protein pH dependency was tested. First freshly purified LKA1 gp49 and its domain were filtered through 10 K Amicon filters and elution buffer was exchanged to PBS at pH range of 1 - 11. After 1 h incubation at room temperature 10µl sample from each preparation was spotted on *P. aeruginosa* PAO1 Krylov lawn. Both, LKA1 gp49 and its domain presented activity in a broad pH range of 1- 11 (Fig. 7.18).

A

B

Figure 7.18: The pH dependency of LKA1 gp49 and its domain. A 10 µl sample of (A) LKA1 gp49 (0.5 mg/ml, 6.22 µM) and (B) LKA1 gp49 domain (0.1 mg/ml, 4.80 µM) treated for 1 h treatment at room temperature with PBS adjusted to pH between 1 and 11, was spotted on *P. aeruginosa* PAO1 lawn. PBS with pH between 1-11 was also spotted and served as a negative control.

7.6. Discussion

The polysaccharide depolymerases have the potential to degrade bacterial biofilm and EPS/LPS, the important factors of the *P.* aeruginosa pathogenesis (Yan *et al.*, 2014). This would allow to disperse biofilm embedded cells or make bacterial cell wall surface more accessible for antibacterial agents, such as antibiotics, detergents, disinfectants, phages or host immune system components. To date polysaccharide depolymerases were only reported in the tailed dsDNA phages from order *Caudovirales* (Sutherland, 1999; Casjens *et al.*, 2011).

In this study we identified and characterized gp49, a pectin lyase and its domain as well as an acylhydrolase domain of gp56, encoded by *P. aeruginosa phiKMVvirus* LKA1 and *Luz7virus* LUZ7, respectively.

The gp49 domain and gp56 domain probably aggregated and their native structure should be further evaluated. Crystal structure of LKA1 gp49 allowed to precisely locate the enzyme active center, which in the case of gp56 could only be predicted based

on protein homology. To our knowledge the depolymerase crystal structure was solved only in the case of the catalytic part of the K1F endosialidase, endoNF (Stummeyer *et al.*, 2005) or gp47, a tail spike protein with endosialidase activity, highly identical in *E. coli* phages K1E and K1-5 (Leiman *et al.*, 2007).

LKA1 gp49 and its domain formed the halo zones on *P. aeruginosa* PAO1 lawn. In case of LUZ7 domain halo formation was absent. Further microscopic analysis showed, that EPS material is keeping bacteria closely in small aggregates, while after protein treatment bacteria were spread out and their EPS slime were reduced. The same observation was presented in several studies (e.g. Cornelissen *et al.*, 2011, 2012; Dubos *et al.*, 1931). Furthermore, with the use of *P. aeruginosa* PAO1 mutants it was elucidated, that all protein samples were able to degrade the B-band O-antigen, while with B-band deficient PAO1 strains the halo zone was not present, as it was also reported for LKA1 phage in A. Cornelissen studies.

The biofilm assay with the use of peg lid plates and CV staining method did not revealed antibiofilm activity of recombinant proteins used in this study. The biofilm degradation analysis should be repeated and discussed in details in the future, preferably with more qualitative and quantitative techniques. In literature, we can find, that Turk *et al.* (2013) and Gutiérrez *et al.* (2015) demonstrated the biofilm-degrading activity of DispersinB and Dpo7, EPS-degrading enzymes active against the staphylococcal biofilm. Biofilm matrix was reduced up to 85% and 90%, respectively.

The stability test revealed that both LKA1 gp49 and its domain are stable in a broad pH range. LKA1 gp49 was also very thermostable and presented the activity even after 1 h incubation at 90°C. Interestingly, its shorter version was less stable and presented activity only up to 70°C. This difference in thermal stability possibly is influence by the absence of full length C-terminal region, responsible for protein trimerization. We hypothesize, that gp49 domain creates a trimeric structure, because it possess partial amino acid sequence from gp49 C-terminus. However, this trimer is less stable and disintegrates after application of high temperatures. It would be interesting to create a new domain, i.e. extended till the end of C-terminal region of gp49 (389 - 769 aa), to assess the influence of C-terminal region of gp49.

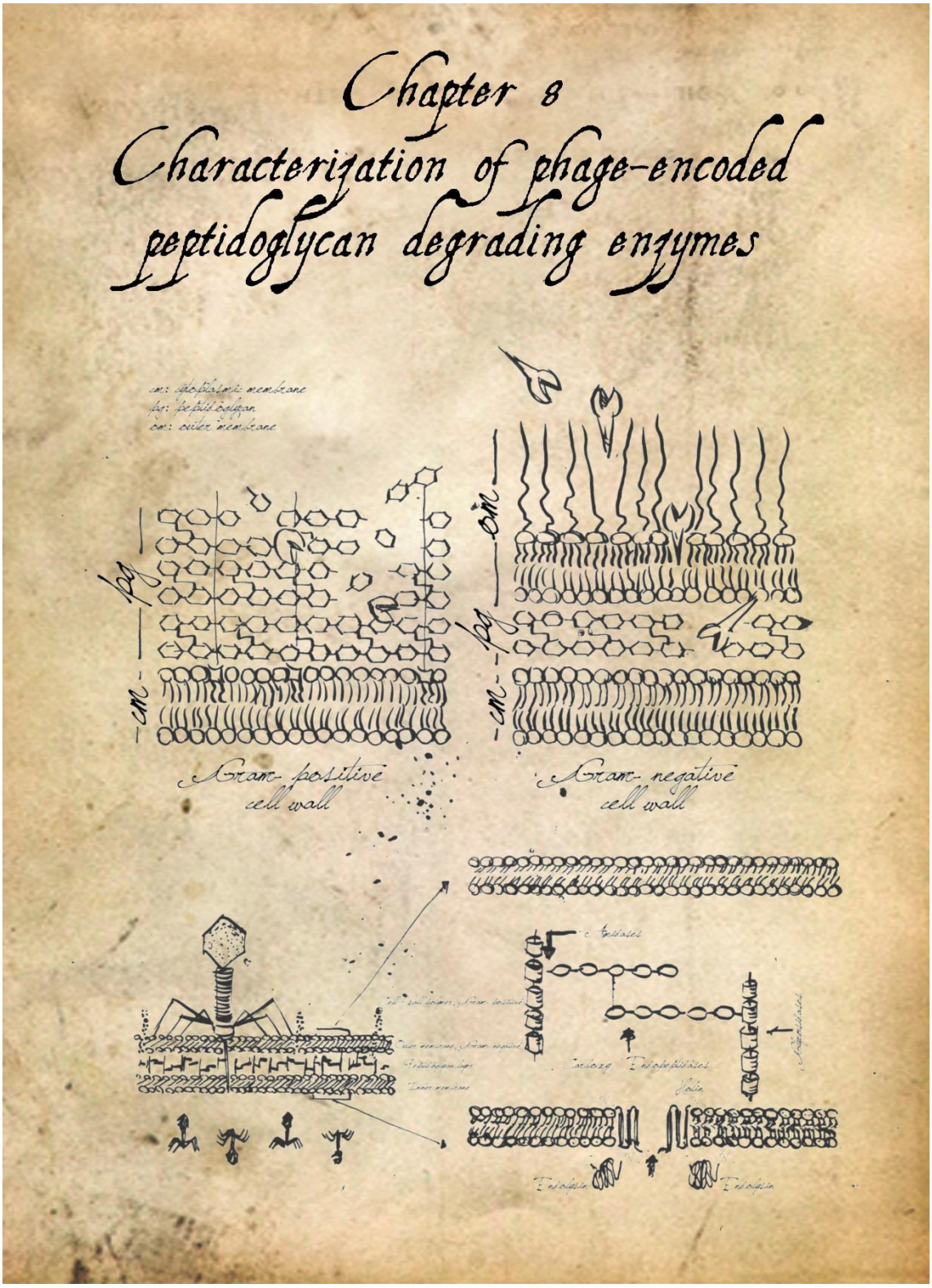
In conclusion, we confirmed that LKA1 gp49 is LPS depolymerase, however its influence on bacterial LPS should be further analyzed in details. Moreover, further analysis should be conducted, focusing also on e.g. protein mode of action and substrate specificity with the use of different *Pseudomonas* biofilm matrix components (i.e. Psl,

Pel or alginate, or pure LPS). The bacterial EPS can also be used to quantify the hydrolyzing activity of domain according to e.g. Chai *et al*. (2014) method, were the purified protein was incubated with EPS substrate for 4 h at 32°C and the reduction of sugars that were released was measured by spectrophotometry and analyzed by HPLC. Moreover, depolymerase domains could be fused with other polysaccharide depolymerase and/or endolysin domains and/or antibacterial peptides with the use of synthetic biology techniques.

Encapsulated bacteria can lead to broad range of severe infections, such as meningitis, septicemia, osteomyelitis, pneumonia, septic arthritis etc. (Moxon *et al.*, 1990). Mushtaq *et al.* (2005) have proven, that systemic infection caused by encapsulated pathogens may be treated with EPS-degrading enzymes. In their study, the low doses of the endosialidase E administrated intraperitoneally to rats at early stage of infection of *E. coli* K1 strain increased the survival rate by removal of bacterial polysaccharide capsule, in comparison to lethal systemic infection in control samples. Moreover, the *per os* administration of tail spike protein P22sTsp reduced also *Salmonella* infections in chickens (Waseh *et al.*, 2010).

We conclude, that polysaccharide depolymerases present an innovative approach, that exploits phage natural weapons. Enzymes could serve as a new antibacterial tool, that enables the access to bacterial cell for commonly used antimicrobials or host immune system. The polysaccharide depolymerases could also be used for the treatment of plant and animal infections (Kim *et al.*, 2000; Mushtaq *et al.*, 2004; Scorpio *et al.*, 2008), eradication or prevention of biofilm formation (Hughes *et al.*, 1998; Verma *et al.*, 2010) or oligosaccharides production derived from polysaccharides (Niemann *et al*., 1977; Altmann *et al*., 1990).

Chapter 8
Characterization of phage-encoded peptidoglycan degrading enzymes



Contribution

- Experiments were conducted by Katarzyna Danis-Włodarczyk and two master students, Michaël Van Ryne and Laura Bukenbergs (2016 master program at KU Leuven Bioscience Engineering, supervisor Katarzyna Danis-Włodarczyk, promoters: Prof. Rob Lavigne and Dr. Lorena Rodriguez-Rubio).
- Time-lapse microscopy was performed in the Centre for Food and Microbial Technology (Prof. Abram Aertsen, KU Leuven, Belgium) with a help of Dr. Sander Govers.
- Additional scientific consultancy was made with Prof. Yves Briers (Ghent University, Belgium) and Dr. Lorena Rodriguez-Rubio (KU Leuven, Belgium), specialists in the field of peptidoglycan degrading enzymes.

8.1. Introduction

Phage PG degrading enzymes (VAPGHs and endolysins) are rapid, highly specific and efficient enzymes, that exert their action at the beginning and at the end of virus reproduction cycle, respectively. VAPGHs operate exogenously, assisting in the viral genetic material injection, while endolysins accumulate in cytoplasm and acts from within, allowing the release of a newly assembled phage progeny (Rodriguez-Rubio *et al.*, 2016). In last decades, several recombinant PG hydrolases have proven to be bactericidal towards Gram-positive bacteria, such as *Staphylococcus*, *Streptococcus* and *Bacillus* strains, when applied *in vitro* or with the use of different animal infection models (Nelson *et al.*, 2001; Loeffler *et al.*, 2001; Schmelcher *et al.*, 2012; Schuch *et al.*, 2002). However, in the case of Gram-negative bacteria, the situation is more complicated, since the OM is a highly effective permeability barrier. To overcome this obstacle, PG degrading enzymes must be delivered in combination with a high hydrostatic pressure (150 - 200 MPa) or OM permeabilizers, such as EDTA, citric acid or malic acid (Briers *et al.*, 2011; Oliveira *et al.*, 2014). Nevertheless, this recombinant phage borne proteins have proven to be highly effective against permeabilized cells of *P. aeruginosa*, *P. putida*, *Burhholderia pseudomallei*, *E. coli*, *S.* Typhimurium, *Acinetobacter baumannii* or *Klebsiella pneumoniae* strains (Briers *et al.*, 2014b; Oliveira *et al.*, 2014; Walmagh *et al*., 2013).

Recently, Briers *et al.* (2014a) presented also another approach focusing on engineered endolysins named ‘Artilysin®s’. These engineered enzymes consists of a recombinant modular endolysin and an antimicrobial peptide, enabling transport through the OM (Briers *et al.*, 2014b). Bacteriolytic phage-borne proteins are currently of high interest as an alternative antimicrobials (Briers *et al.*, 2014). Importantly, to date

there was no case of resistance occurrence *in vitro* and *in vivo* (Schuch *et al.*, 2002; Rodriguez-Rubio *et al.*, 2013; Loeffler *et al.*, 2003; Pastagia *et al.*, 2011, O'Flaherty *et al.*, 2009).

In this chapter we focus on PG degrading enzymes of *P. aeruginosa* phages presented in chapters 4 - 6. With the use of bioinformatics analysis six putative PG hydrolases genes were identified. Based on their homology with other phage proteins, the theoretical secondary and tertiary structure was elucidated. Further, the methodology of recombinant protein expression and purification was optimized, allowing to achieve pure and active phage enzymes. In the final part, bactericidal activity and stability of phage enzymes was examined.

8.2. Bioinformatics analysis, a search for genes encoding PG degrading enzymes

The present dissertation started with the analysis of phage genomes (chapters 4 - 6) and identification of the promising phage ORFs encoding PG degrading enzymes, with the use of BLASTP, HHpred, HMMER, Phyre2 and InterPro bioinformatics tools (Altschul *et al.*, 1990; Remmert *et al.*, 2011; Finn *et al.*, 2011; Kelley *et al.*, 2015; Jones *et al.*, 2014). Table 8.1. presents the list of selected phage-encoded proteins and their basic features. All chosen proteins exhibit lower than 25% similarity to each other, except two proteins, KT28 gp49 and KTN6 gp46 with only 4 amino acids difference in the amino acid sequence (98.2% identity).

Table. 8.1: List of PG degrading enzymes candidates chosen for study.

Protein (NCBI code)	Origin	Nucleotide length	Amino acids length	Theoretical pI	Ext. coeff. (M^{-1} cm^{-1})	Predicted structure	Predicted catalytic specificity	Phage protein homologue (BLASTP, NCBI code, e-value)
KT28 gp49 (AKJ71487) putative endolysin	*P. aeruginosa* KT28 phage (*Pbunavirus*)	663	220 (24.14 kDa)	8.76	39545	globular	chitinase, glycoside hydrolase 19 family, lysozyme-like domain,	putative lysozyme gp48 of *P. aeruginosa* phage SN (YP_002418854.1, e-value 7e-163)
KT28 gp41 (AKJ71479.1) putative VAPGH		2577	942 (94.23 kDa)	8.20	58790	globular	lytic transglycosidase	putative transglycosidase gp40 of *P. aeruginosa* phage SN (YP_002418846.1, e-value 0.0)
KT28 gp41 domain (463-605 aa) putative VAPGH		432	144 (16.22 kDa)	9.36	17420			
KTN6 gp46 (AKJ71578.1) putative endolysin	*P. aeruginosa* KTN6 phage (*Pbunavirus*)	663	220 (24.17 kDa)	8.76	39545	globular	chitinase, glycoside hydrolase 19 family, lysozyme-like domain	putative endolysin gp45 of *P. aeruginosa* phage KPP12 (YP_007238200.1, e-value 6e-163
PA5oct gp214 putative VAPGH	*P. aeruginosa* PA5oct phage (novel giant)	1434	447 (51.08 kDa)	5.94	52495	globular	tail associated lysozyme	baseplate hub subunit and tail lysozyme gp234 of *Cronobacter* phage GAP32 (YP_006987339.1, e-value 7e-48)
PA5oct gp250 putative endolysin		516	172 (18.54 kDa)	9.26	11710	globular	glycoside hydrolase 19 family, lysozyme-like domain	-
KTN4 gp48 (ANM44806.1) putative endolysin	*P. aeruginosa* KTN4 phage (*Phikzvirus*)	807	271 (29.70 kDa)	3.82	18450	globular	N-acetylmuramoyl-L-alanine amidase	gp35 of *P. aeruginosa* phage φKZ (NP_803601.1, e-value 0.0)

Abbreviations: (pI) a isoelectric point, (Ext. coeff.) a extinction coefficient parameter. The protein size in kDa was calculated with ExPASy ProtParam, putative catalytic specificity is established based on homology analysis performed with the use of BLAST P, HMMER, HHPRED a PHYRE2.

Based on sequence homology four phage proteins, KT28 gp49, KTN6 gp46, PA5oct gp250 and KTN4 gp48, are predicted to be globular endolysins. KT28 gp41 and PA5oct gp214 are homologues to virion- associated lytic proteins, thus they are potential VAPGHs.

KT28 gp49

Pfam analysis revealed, that gp49 of *Pbunavirus* KT28 (chapter 4) consist a glycoside hydrolase family 19 chitinase class I domain [locus in full length protein: 103 - 217 amino acids (aa)]. Prediction by the InterPro Scan showed presence of two potential locus of lysozyme-like domains (13 - 76 aa and 108 - 216 aa) and one of glycoside hydrolase family 19 (122 - 177 aa). The glycoside hydrolases are enzymes, that the β-1,4 linkages between the GluNAc and MurNAc sugars of PG (Wolhlkonig *et al.*, 2010). It is very broad group of enzymes, consisting >100 different families (http://www.cazy.org/GH1.html; http://enzyme.expasy.org/ EC/3.2.1.-). The glycoside hydrolase family 19 comprises enzymes with chitinase activity, that catalyse the hydrolysis of the β-1,4-N-acetyl-D-glucosamine linkages in chitin polymers.

KTN6 gp46

Gp46 is a protein of *Pbunavirus* KTN6 (chapter 4). It presents 98.18% of amino acid sequence identity to KT28 gp49 (Clustal Omega) with 4 point amino acids alterations (Fig. 8.1).

```
KTN6_gp46  MKITKDILITGTGCTTDRAIKWLDDVQAAMDKFHIESPRAIAAYLANIGVESGGLVSLVE
KT28_gp49  MKITKDVLITGTGCTTDRAIKWLDDVQTAMDKFHIESPRAIAAYLANIGVESGGLVSLVE

KTN6_gp46  NLNYSAQGLANTWPRRYAVDPRVRPYVPNALANRLARNPVAIANNVYADRMGNGCEQDGD
KT28_gp49  NLNYSAQGLANTWPRRYAVDPRVRPYAPNALANRLARNPVAIANNVYADRMGNGCEQDGD

KTN6_gp46  GWKYRGRGLIQLTGKSNYSLFAEDSGMDVLEKPELLETPAGASMSSAWFFWRNRCIPMAE
KT28_gp49  GWKYRGRGLIQLTGKSNYSLFAEDSGMDVLEKPELLETPAGASMSSAWFFWRNRCIPMAE

KTN6_gp46  SNNFSMVVKTINGAAPNDANHGQLRINRYMKTIAAINQGS
KT28_gp49  SNNFSMVVKTINGAAPNDANHGQLRINRYLKTIAAINQGS
```

Figure 8.1: The amino acid sequence comparison between KT28 gp49 and KTN6 gp46 (Clustal Omega v. 1.2.1). Sequence point alterations are marked gray.

Amino acid substitution in position 7 [isoleucine (I, Ile) changed to valine (V, Val)] and 210 [methionine (M, Met) changed to leucine (L, Leu)] seems to be neutral. The former pair have 1.67 and 1.87 preference to β-strands, respectively, while the latter have 1.30 and 1.34 preference to α-helix, respectively (according the Chou-Fasman method, 1974). A different situation exists in a) position 28, where alanine (A, Ala)

is changed to threonine (T, Thr). First is a strong former of a helix and a weak former of a strand with 1.41 preference to α-helix, second is a former of a strand and indifferent towards helix with 1.45 preference to β-strands; b) position 87, where valine (V, Val) is change to alanine (A, Ala). Val is a strong former of strand and a former of helix, with 1.87 preference to β-strands, while Ala is strong helix former and a weak former of strand, with 1.41 preference to α-helix. Clearly substituted amino acids have different properties, however, without protein crystal structure it is difficult to assess the impact on protein structure and folding, that could lead to differences in antibacterial activity, substrate specificity or stability. Nevertheless, as it will be shown later, enzymes had different bacteriolytic activity.

The HMMER analysis also elucidated presence of glycoside hydrolase family 19 domain (96 - 219 aa). The InterPro Scan showed three possible domains with almost identical locus as it was presented in case of KT28 gp49: lysozyme-like domains (12 - 76 aa and 108 - 216 aa) and glycoside hydrolase family 19 (121 - 177 aa).

PA5oct gp250

Gp250 is a protein of a novel giant phage PA5oct (chapter 6). It comprise a lysozyme-like domain [pfam (42 - 161 aa) and InterPro Scan (11 - 170 aa) prediction]. Lysozyme has activity of muramidase (N-acetylmuramide glycanhydrolase) and catalyze hydrolysis of 1,4-β-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in a bacterial peptidoglycan. It also belongs to group of glycoside hydrolases (Yoshimura *et al.*, 1988).

KTN4 gp48

Gp48 is a protein of *Phikzvirus* KTN4 (chapter 5). The Pfam analysis showed, that its amino acid sequence is homologues to a peptidase and glycoside hydrolase family 18, however it was not possible to predict domain locus. According to prediction this protein can exhibit N-acetylmuramoyl-L-alanine amidase activity and hydrolyse the amid bond that anchors the L-alanine moiety of the pentapeptide chain to the D- lactyl group of the MurNAc sugar in the glycan backbone (Cheng *et al.*, 1994).

KT28 gp41

Gp41 is potential VAPGH of *Pbunavirus* KT28. The pfam and InterPro Scan analysis revealed the presence of a transglycosylase SLT (soluble) domain (pfam 489 - 593 aa; InterPro Scan 480 - 594 aa) (Fig. 8.4). The lytic transglycosylases degrade murein via cleavage of the β-1,4-glycosidic bond between GluNAc and MurNAc in bacterial PG, followed by the formation of a 1,6-anhydromuramic acid terminal residue (Bienkowska-Szewczyk *et al.*, 1980). This protein was very difficult to handle during cloning and expression process, as it will be further discuss in section 8.2.2. That is why we decided to focus on its active domain (Fig. 8.2). Additional margin of amino acids was added on C- and N-termini of predicted domain, to avoid cuts in secondary structure elements, i.e. helices and strands, that could have influence on protein folding and function. Finally, domain was defined between 463 and 605 amino acid residue of full length protein.

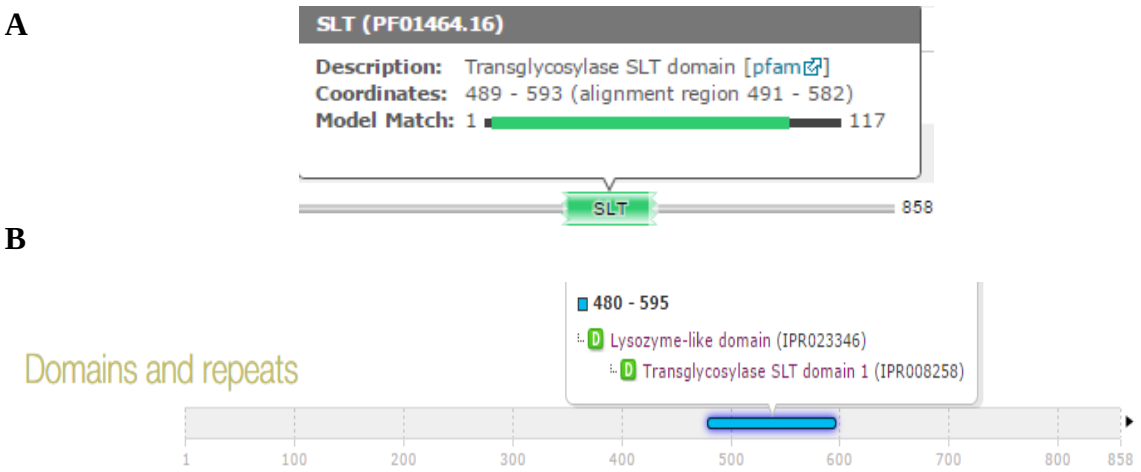


Figure 8.2: The protein homology analysis of KT28 gp41 with prediction of functional domains. (A) analysis performed by HMMER, prediction of transglycosylase SLT domain (489 - 593 aa locus). (B) analysis performed by InterPro Scan, elucidation of lysozyme- like domain or transglycoslylase SLT domain 1, located between 480 and 595 amino acid.

PA5oct gp214

The potential VAPGH gp214 of the giant phage PA5oct consist of a phage tail associated lysozyme-like domain [(42 - 161 aa) pfam, (11 - 170 aa) InterPro Scan].

All protein activity was assigned only based on available homology in databases. For proper catalytic specificity elucidation a detail biochemical analysis has to be introduced.

Further potential protein secondary and tertiary structure was predicted and analyzed for all proteins with the use of Phyre2 (Kelley *et al.*, 2015), PyMol Molecular Graphics System v. 1.8 Schrödinger (Fig. 8.3, Table 8.2.). This analysis was also essential for KT28 gp41 domain definition.

Table 8.2: Prediction of the protein secondary structure composition with elucidation of ligand binding site.

Protein	Secondary structure composition	The closest protein structure homologue	Amino acid composition of ligand binding side
KT28 gp49	12 α-helices, 1 β-strand, 8% disordered structure	endolysin spn1a from S. Typhimurium with glycosidase activity through the catalytic dyad (96% and 97% coverage, respectively; PDB code 4OK7, Park *et al.*, 2014)	130 ILE, 131 GLN, 132 LEU, 169 PHE, **172 ARG** 191 ILE
KTN6 gp46			130 ILE, 131 GLN, 169 PHE, 191 ILE **192 ASN**
KT28 gp41	41 α-helices, 2 β-strand, 31% disordered structure	complex of the 70 kDa soluble lytic transglycosylase with bulgecin A, an O-sulfonated glycopeptide with N-acetyl-D-glucosamine binding side (71% and 94% coverage, respectively; PDB code 1SLY, Thunnissen *et al.*, 1995)	-
KT28 gp41 domain	8 α-helices, 11% disordered structure		69 GLN, 70 LEU, 71 MET, 74 ILE, 96 LEU, 114 TYR, 115 TYR, 116 HIS, 117 GLY, 118 GLY, 119 LEU, 120 ASP, 123 ASN
PA5oct gp214	3 α-helices, 41 β-strand, 52% disordered structure	pa0091 vgrg1 protein, the central spike of the type vi2 secretion system (84% coverage, PDB code 4MTK, Sycheva *et al.*, to be published) as well as T4 bacteriophage needle proteins gp27 and gp5 (25% coverage, PDB code 2P5Z, Leiman *et al.*, 2009) *	-
PA5oct gp250	7 α-helices, 3 β-strand, 17% disordered structure	R21 endolysin from coliphage 21 with N-terminal signal-anchor-release (SAR) domain (98% coverage, PDB code 3HDE, Sun *et al.*, 2009)	30 GLY, 48 GLY, 101 TYR, 102 ASN, 105 PRO
KTN4 gp48	5 α-helices, 3 β-strand, 65% disordered structure	rfnbpa in compex with fibrinogen 2 gamma chain c-terminal peptide (22% coverage, PDB code 4B60, Stemberk *et al.*, 2014) *	-

(-) no prediction due to low homology score. Bolded amino acids represent the difference in composition of predicted binding side between highly similar enzymes, KT28 gp49 and KTN6 gp46. All secondary structure components are left- handed.

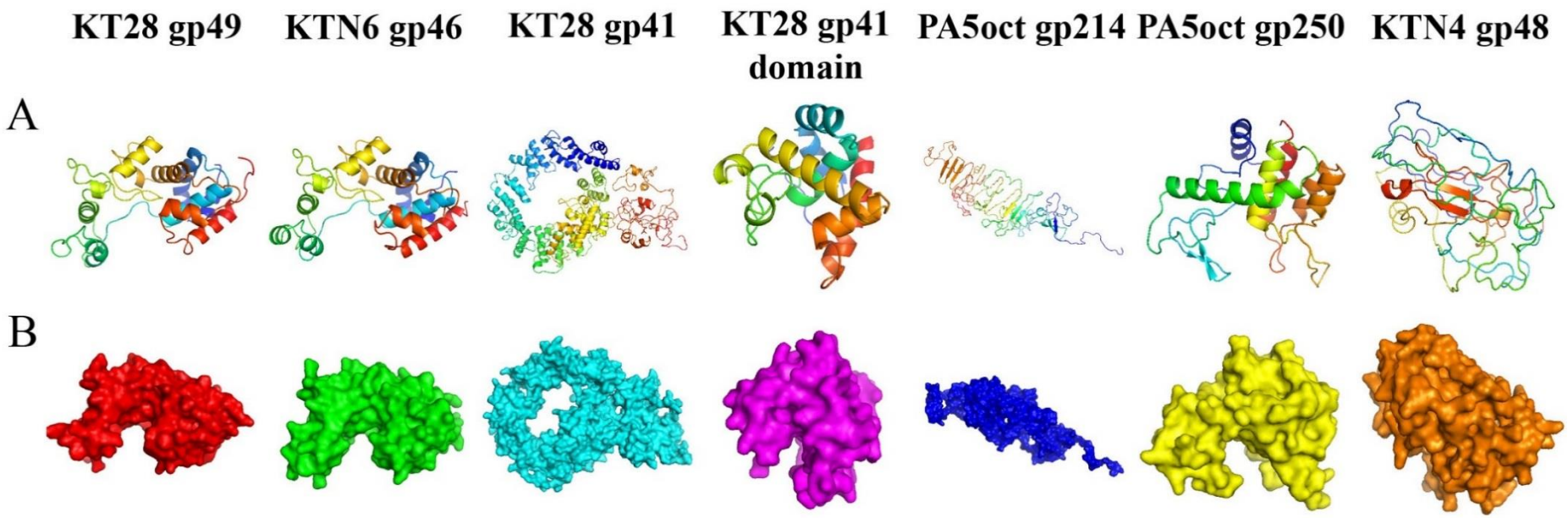


Figure 8.3: Prediction of tertiary protein structure based on a structure homology. (A) protein structure represented by left handed α-helices, β-strands, loops and turns as a cartoon, (B) protein structure presented as a surface. Protein structures are not presented in a scale.

Furthermore, proteins ligand binding site was predicted based on structure homology with the use of 3DLigandSite –Ligand binding site prediction server (Wass *et al*., 2010) (Fig. 8.4). In the case of KT28 gp41, PA5oct gp214 and KTN4 gp48 ligand binding side identification was not possible due to low homology. Selected amino acids were located in the center of protein structure and could serve in the future as a target for side-directed mutagenesis experiments in order to modify enzyme specificity.

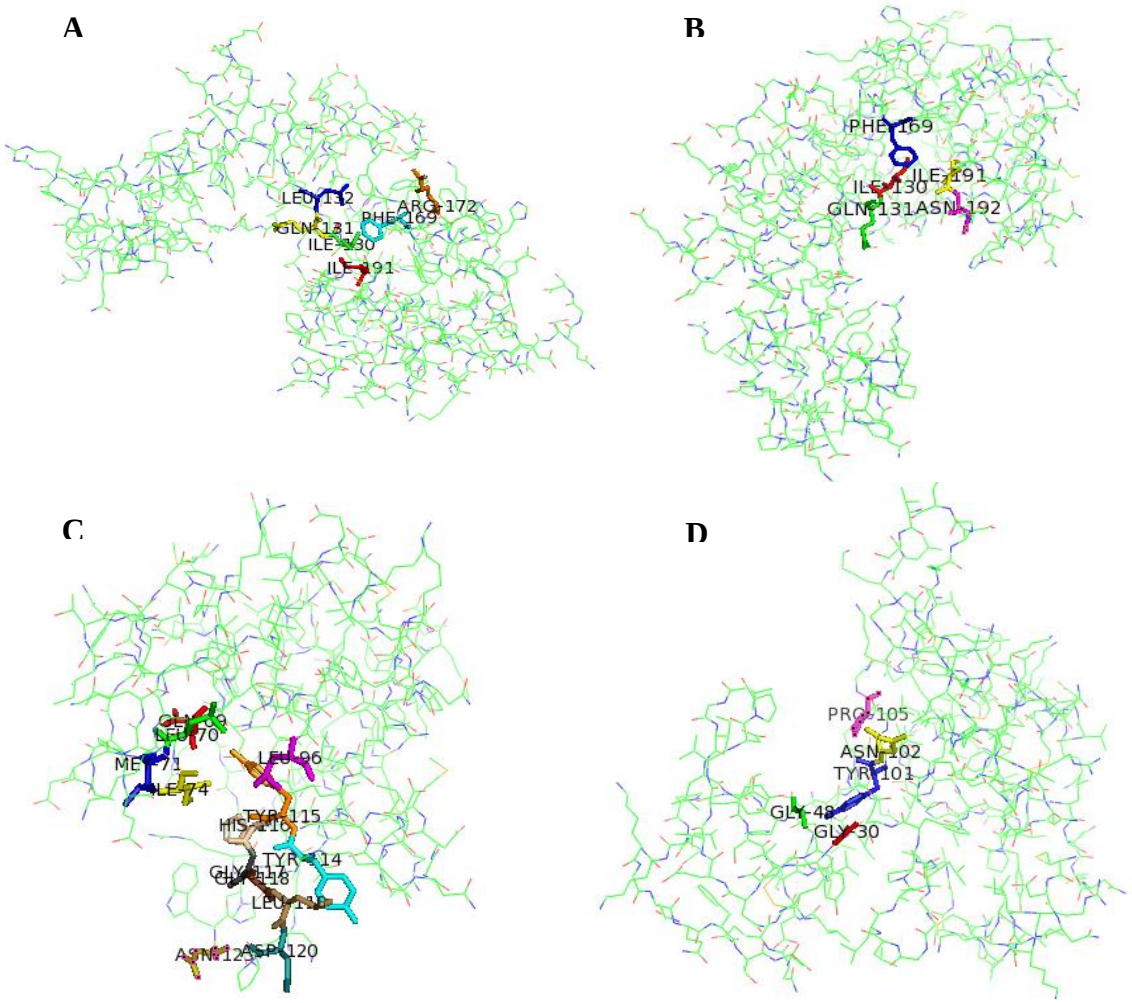


Figure 8.4: The predicted ligand binding sides of (A) KT28 gp49, (B) KTN6 gp46, (C) KT28 gp41 domain and (D) PA5oct gp250. Crucial amino acids from Table 8.2 are represented as colorful sticks. The rest of protein structure is shown as lines. Structures were modelled by PyMol v. 1.8 Schrödinger. Structures are not presented in scale.

All above analysis can serve as a basis for potential future protein engineering experiments or understanding mode of action and enzyme kinetics. However, more in depth analysis should be performed, preferably based on protein crystallography, due to low homology score and coverage in case of some proteins.

8.3. Cloning, recombinant expression and purification

The full-length ORFs encoding six chosen PG degrading enzymes (Table 8.1) were isolated and amplified by Phusion® High-Fidelity DNA polymerase with low error rate. Next, all PCR products were introduced into pEXP5-CT/TOPO® vector. Produced proteins carried a 6x His-tag on C-termini, important for further protein purification steps. Cloning was successful for KT28 gp49, KTN6 gp46, PA5oct gp214 and KTN4 gp48. However, in the case of PA5oct gp250 and KT28 gp41 constructs had several point mutations, especially on C-termini. Despite several repetitions of gene amplification and cloning, no correct constructs could be obtained. That is why PCR products were further successfully introduced into another vector, the pEXP5-NT/TOPO® with 6x His-tag on N-termini.

Next, protein expression optimization was conducted in 100 ml cultures with the use of different *E. coli* chemicompetent strains, medium, expression temperature and inducer concentration (Table 8.3) in order to achieve the most abundant protein expression in soluble fraction. First different conditions were tested with KT28 gp49 and KTN6 gp46, to reduce the number of variables for the rest of proteins. Unfortunately, in the case of full length KT28 gp41 expression failed and protein was continuously present in insoluble fraction or expression did not occurred, regardless of a variety of expression conditions. To overcome this obstacle KT28 gp41 domain was identified and successfully cloned into pEXP5-NT/TOPO® vector (Table 8.3). The solubility of proteins was checked with the use of SDS-PAGE and Western Blot methods (S Fig. 8.1).

Based on a broad range expression optimization, for this particular set of proteins it can be concluded, that :1) the most suitable *E. coli* expression strains were Lemo21 (DE3) and BL21 (DE3) pLysS, both under control of T7 promoter; 2) the most universal for protein expression was 2xTY medium, 3) expression at 16°C allowed bacteria to produce proteins slower with higher abundance in soluble fraction, 4) the protein yield was the same regardless of inducer concentration used.

Table. 8.3: Protein expression optimization process for all chosen PG degrading enzymes.

Type	Name	KT28 gp49	KTN6 gp46	KT28 gp41	KT28 gp41 domain	PA5oct gp214	PA5oct gp250	KTN4 gp48
vector	pEXP5-CT/TOPO	√	√			√		√
	pEXP5-NT/TOPO			√	√		√	
expression strain	Lemo21 (DE3)	√	√	√	√	√	√	√
	BL21 CodonPlus (DE3)			√	√	√	√	√
	BL21 (DE3) pLysS			√			√	√
	BL21 AI			√			√	
	Arctic Express (DE3)			√			√	
medium	LB	√	√	√	√	√	√	√
	2xTY	√	√	√	√	√	√	√
	Modified LB	√	√	√	√	√		
	TB	√	√	√	√			
	Autoinduction	√	√	√	√			
temperature post induction	37°C (4.5 h)	√	√		√	√		
	16°C (overnight)	√	√	√	√	√	√	√
L-rhamnose	2 mM	√	√					
	1 mM	√	√					
	0.5 mM	√	√	√	√	√	√	√
	0.1 mM	√	√					
IPTG	0.1 mM	√	√	√		√		√
	0.4 mM	√	√	√	√		√	
final conditions		Lemo21 (DE3) LB medium 0.4 mM IPTG 0.5 mM L-rhamnose 16°C (overnight expression)	Lemo21 (DE3) 2xTY medium 0.4 mM IPTG 0.5 mM L-rhamnose 16°C (overnight expression)	Lack of suitable conditions	Lemo21 (DE3) 2xTY medium 0.4 mM IPTG 0.5 mM L-rhamnose 16°C (overnight expression)	CodonPlus™ 2xTY medium 0.1 mM IPTG 16°C (overnight expression)	Lemo21 (DE3) 2xTY medium 0.4 mM IPTG 0.5 mM L-rhamnose 16°C (overnight expression)	BL21 pLysS 2xTY medium 0.1 mM IPTG 16°C (overnight expression)

Gray color marks conditions were expression was successful and protein appeared in soluble fraction.

The next step was a large scale purification performed with the use of ÄKTA FPLC system in combination with HisTrap^{TM} HP 1 ml columns or gravity-flow chromatography at 4°C with a Ni^{2+}-NTA resin (Fig. 8.5). Purity of all proteins was further increased by filtration through 10 K filter units. For KTN6 gp46 and PA5oct gp214 it was necessary to apply additional purification step, the size exclusion chromatography with the use of the HiLoad 16/600 Superdex 200 Prep grade gel filtration column and the HiLoad 16/600 Superdex 75 Prep grade gel filtration column, respectively (S Fig. 8.1). Highly pure proteins were kept in elution buffer at 4°C with precipitation inspections. In case of KT28 gp49 and KTN6 gp46 precipitation occurred after few days. To improve the stability, all proteins were kept further in 5mM HEPES buffer with 30% glycerol, proven in several studies to be a suitable buffer for endolysins as well as proteins in a broader sense (Briers *et al.*, 2014a, 2011; Oliveira *et al.*, 2016).

Based on size exclusion chromatography, the native state of PA5oct gp214 and KTN6 gp46 was estimated (52.1 kDa and 20.98 kDa, repectively). This suggest, that both proteins occur as a monomeric structures, however further analysis should be conducted, e.g. crystallography.

Table 8.4 presents the protein yields of different PG degrading enzymes with the use of three different protein purification techniques. The highest yield was achieved for KTN6 gp46 (3.37 g/750 ml culture). The lowest yields and the most complicated expression and purification were in case of KT28 gp41 and KTN4 gp48. Preferably other expression systems should be introduced, such as expression in yeast system, to which proteins should not be toxic. For PA5oct gp250 and gp214 protein purity increased with the use of 10 K filter units, however further purification steps and optimization should be introduced.

Table 8.4. The average yields of purified proteins. Measurements were performed on a Nanodrop at wavelength of 280nm with specific parameters of molecular weight and extinction coefficient (Expasy Protparam).

Protein	Average protein yield (mg) after gravity affinity purification	Total protein yield (mg) after FPLC affinity purification*	Total protein yield (mg) after size exclusion purification*
	750 ml culture	750 ml culture	750 ml culture
KT28 gp49	4.84	3375	N/A
KTN6 gp46	5.77	2578	1100
KT28 gp41 domain	1.0	3.50	N/A
PA5oct gp214	1.54	N/A	0.50
PA5oct gp250	1.62	0.67	N/A
KTN4 gp48	0.50	N/A	N/A

(*) in 0.5 ml fractions, (FPLC) Fast Protein Liquid Chromatography

8.4. Assessment of PG degradation activity

To demonstrate the ability of six purified phage enzymes to degrade the PG layer, several techniques were applied.

Zymography with crude peptidoglycan

This technique was used to identify the hydrolytic activity of pure proteins towards crude PG, isolated from *P. aeruginosa* PAO1 Krylov. The positive activity was visualized by appearance of transparent bands on 12% SDS-PAGE gel containing crude PG. Results for six phage PG hydrolases were positive (Fig. 8.5). However, it has to be stressed, that PA5oct gp214 presented very weak band, which was hard to capture on a photo. In order to assess if phage PG hydrolases were associated with viral structural proteins, whole phage proteomes were also loaded on zymograms. Only in the case of PA5oct gp214 a very small transparent band appeared, corresponding to the band from PA5oct proteome (Fig. 8.5).

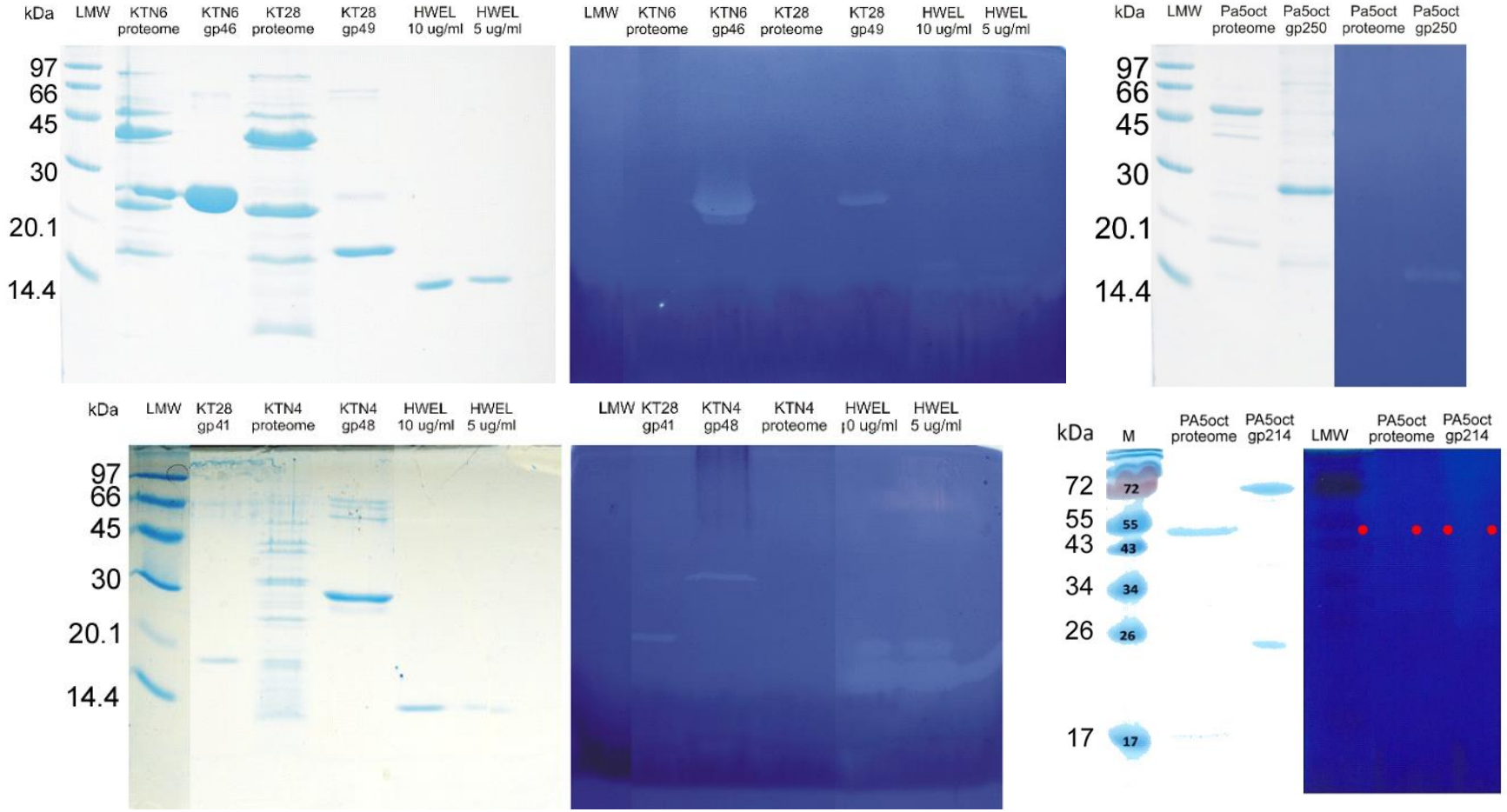


Figure 8.5: Zymograms of six purified phage PG hydrolases. A 10 µl/well of KT28 gp49 (6.88 mg/ml, 24.14 kDa), KT28 gp41d (0.41 mg/ml, 16.22 kDa), KTN6 gp46 (6.08 mg/ml, 24.14 kDa), PA5oct gp214 (1.43 mg/ml, 51.08 kDa), PA5oct gp250 (0.86 mg/ml, 18.54 kDa) or KTN4 gp48 (0.44 mg/ml, 29.70 kDa) in 5 mM HEPES/NaOH buffer (pH 7.4) was loaded on a 12% acrylamide SDS-PAGE gels containing crude PG of *P. aeruginosa* PAO1 Krylov strain. To assess the size of bands indicating PG degradation (light blue bands), identical standard 12% SDS-PAGE gels were performed in the same conditions and stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific). As a positive control served a hen egg-white lysozyme (HEWL) in 10 µg/ml, 5 µg/ml concentrations. (LMW) Low Molecular Weight marker (GE Healthcare), (M) PageRuler Prestained Protein Ladder (Thermo Scientific). [edited images]

Antibacterial assay

Degradation of PG ultimately leads to a killing effect. To demonstrate the bactericidal activity of the phage enzymes, an antibacterial assay was performed in triplicate with the use of *P. aeruginosa* PAO1 Pirnay cells and freshly purified proteins (final concentration of 2.5 μM) kept in 5 mM HEPES/NaOH buffer (pH 7.4) (Fig. 8.6).

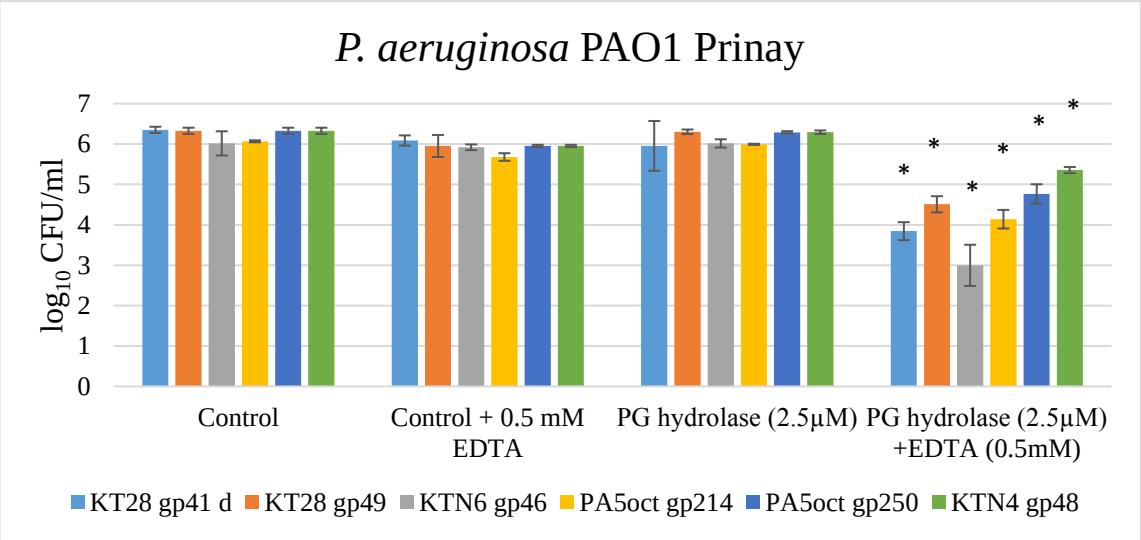


Figure 8.6: Assessment of antibacterial activity of 6 phage PG hydrolases on *P. aeruginosa* PAO1 Prinay strain. (Blue) KT28 gp41 domain, (orange) KT28 gp49, (gray) KTN6 gp46, (yellow) PA5oct gp214, (dark blue) Pa5oct gp250, (green) KTN4 gp48. (Control) *P. aeruginosa* PAO1 Prinay strain mixed with 5 mM HEPES/NaOH buffer; (Control + EDTA) *P. aeruginosa* PAO1 Prinay strain mixed with 5 mM HEPES/NaOH buffer 0.5 mM EDTA; (PG hydrolase) *P. aeruginosa* PAO1 Prinay strain mixed with PG hydrolase in 5 mM HEPES/NaOH buffer (2.5 μM final concentration); (PG hydrolase + EDTA) *P. aeruginosa* PAO1 Prinay strain mixed with PG hydrolase in 5 mM HEPES/NaOH buffer (2.5 μM final concentration) and 0.5 mM EDTA. Columns represent average $\log_{10}$ (CFU/mL) from three independent experiments with error bars indicating the standard deviation. (*) significant log reduction.

Population of *P. aeruginosa* PAO1 Prinay strain was significantly reduced ($p < 0.05$) in the presence of 0.5 mM EDTA in combination with 2.5 μM of KT28 gp41 domain (2.5 ± 0.2 log reduction), KT28 gp49 (1.8 ± 0.23 log reduction), KTN6 gp46 (3.0 ± 0.3 log reduction), PA5oct gp214 (1.9 ± 0.48 log reduction), PA5oct gp250 (1.6 log ± 0.23 reduction), and KTN4 gp48 (1.0 ± 0.5 log reduction). KTN6 gp46 exhibited the strongest bactericidal activity, while KTN4 gp48 the weakest influence. Enzyme concentration used in the study was chosen in order to compare the presented results with the activity of other endolysisns, belonging to the collection of the Laboratory of Gene Technology (KU Leuven) and previously described by Prof. Y. Briers and Dr. L. Rodrigues-Rubio. Preliminary tests with undiluted concentrations of KT28gp49 (14.5 μM) and KTN6 gp46 (20.9 μM) in the presence of 0.5 mM EDTA resulted

in 5.99 log reduction and no survival of *P. aeruginosa* PA01 Pirnay cells, respectively. In the absence of EDTA, no significant antibacterial activity was observed. This justifies the need for an OM permeabilizing agents to promote the access of the protein to the PG substrate. EDTA binds the stabilizing divalent cations in the LPS layer, thereby destabilizing the OM. Application of EDTA alone had no influence on CFUs numbers.

Host range

To verify whether PG hydrolases also demonstrate antibacterial activity towards *E. coli* APEC Ch2 strain and *S.* Enteritidis ATCC13076 strain, an antibacterial assay was performed with freshly purified proteins in 5 mM HEPES/NaOH buffer (pH 7.4) (final concentration 2.5 μM). However, no significant log reduction was observed in the treatment of 0.5 mM EDTA in combination with 2.5 μM of all tested proteins. This can be explained by the lower degree of phosphorylation of the LPS molecules and thus a lower amount of stabilizing divalent cations in the OM of *E. coli* and *Salmonella,* which makes their OM less prone to the action of EDTA. This phenomenon was also observed by other researches (Oliveira *et al.*, 2014 and 2016, Walmagh *et al.*, 2013). Based on Oliveira *et al.* (2014) study other two OM permeabilizers were introduced, 2 mM citric acid or 5 mM malic acid, however results remain the same.

Since KTN6 gp46 exhibited the highest activity, further analysis was focused only on this protein.

Quantitative assessment of enzymatic activity

To quantified the hydrolytic activity of KTN6 gp46, the drop in turbidity of an OM-permeabilized *P. aeruginosa* Pirnay suspension was measured with small intervals after addition of the purified recombinant enzyme (Fig. 8.7). The general formula to express enzymatic activity (activity units/mg) was previously defined by Prof. Y. Briers (Briers *et al.*, 2007).

$$\text{Activity } (^{\text{units}}/_{\text{mg}}) = \frac{\Delta \text{OD } 655\text{nm}/_{(\text{min.mg})}}{0.001} \times \frac{\text{reaction volume } (\mu\text{l})}{1000\ \mu l}$$

First different linear regressions for an increasing data (n = number of measurements in time, starting from n = 5, 6, 7…) were calculated. Next, the corresponding determination coefficient (R^2) indicated the degree of linear relation between optical density and time and it is a measure of how well the linear regression

represents the selected data set. The R^2 will maximize, as more data of the linear regression are included, however it will decrease beyond the linear region. When R^2 was maximized, the slope of the linear regression was calculated and presented as a drop in $OD_{655\ nm}$ per minute. The reliability of the slope will rise with increasing number of measurements with the linear region. Enzyme activity is expressed in units/mg. Excel spreadsheet created by Prof. Y. Briers allowed to increase efficiency in a data statistically processing (Briers *et al.*, 2007). Endolysin KZ144 (0.95 mg/ml) was used as a positive control and 5 mM HEPES/NaOH (pH 7.4) as a negative control.

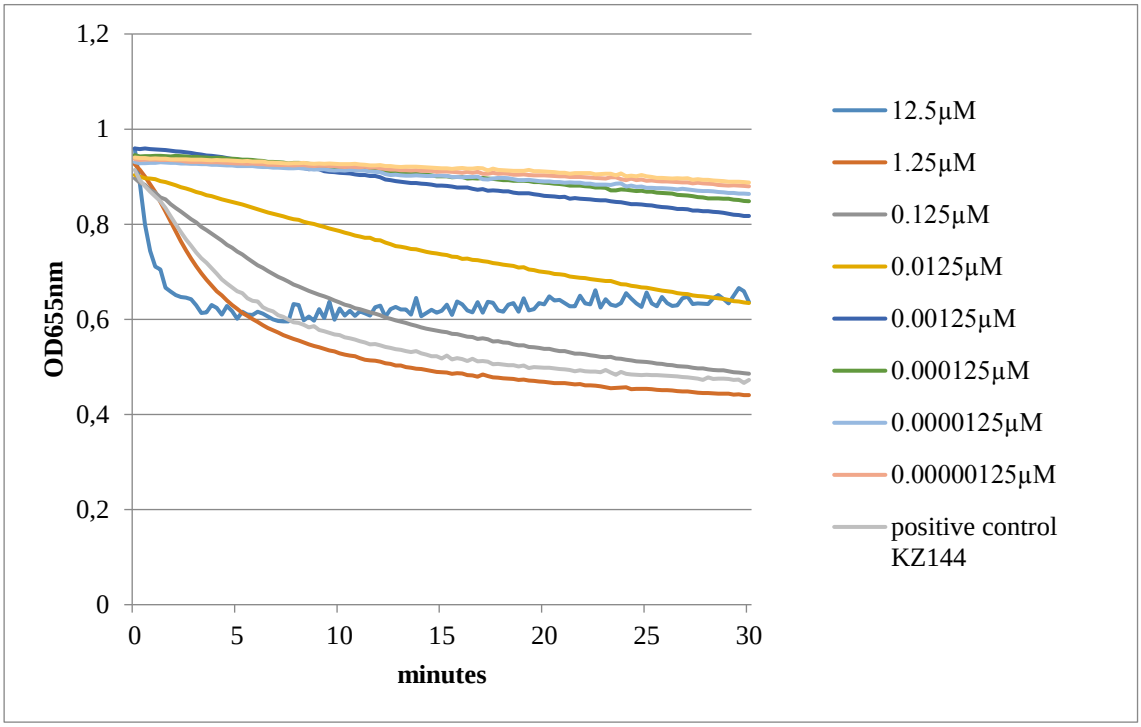


Figure. 8.7: Enzymatic activity of different concentrations of KTN6 gp46 towards *P. aeruginosa* PAO1 Pirnay substrate. Recombinant protein KTN6 gp46 and *P. aeruginosa* substrate were suspended in 5 mM HEPES/NaOH (pH 7.4). As a negative control served 5 mM HEPES/NaOH buffer (pH 7.4) alone and as a positive control endolysin KZ144 (0.95 mg/ml), also in 5 mM HEPES/NaOH buffer (pH 7.4).

Figure 8.8 shows the saturation curve obtained from three independent replicates. Each data point is an average with its standard deviation. For the lower concentrations a linear trend line with a R^2 of 97.3% could be obtained leading to the calculation of muralytic activity in units/µM. The ΔOD_{655nm}/min value increased proportionally to the amount of enzyme added. From 11129 units/µM of enzyme the curve was saturated, which indicates the maximum activity.

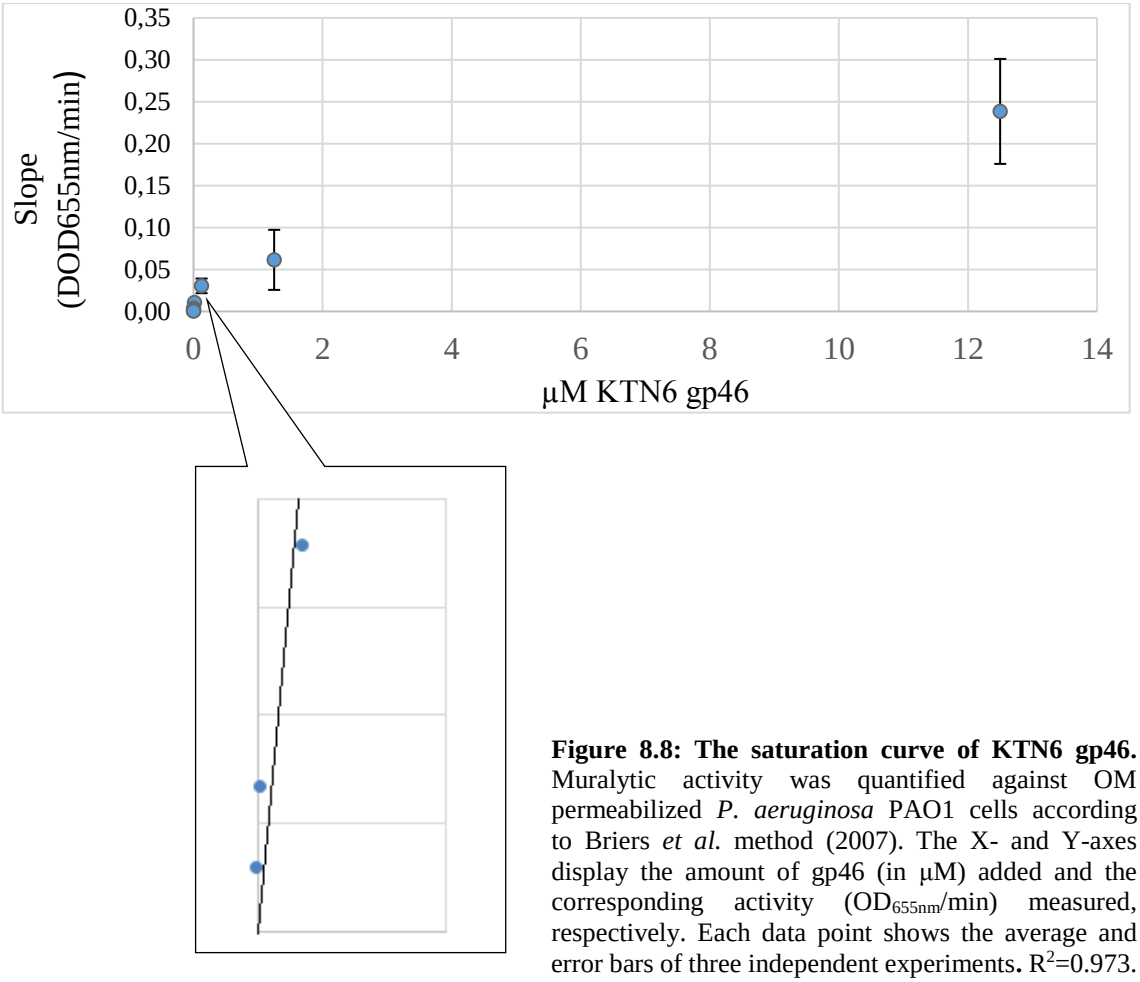


Figure 8.8: The saturation curve of KTN6 gp46. Muralytic activity was quantified against OM permeabilized *P. aeruginosa* PAO1 cells according to Briers *et al.* method (2007). The X- and Y-axes display the amount of gp46 (in µM) added and the corresponding activity (OD_{655nm}/min) measured, respectively. Each data point shows the average and error bars of three independent experiments**.** R^2=0.973.

Visualizing individual cell lysis using time-lapse microscopy

To visualize the mode of action of KTN6 gp46 a time-lapse microscopy was applied with the use of a phase-contrast microscope (Prof. A. Aertsen, Department of Microbial and Molecular Systems, KU Leuven). With the highest concentrations of KTN6 gp46 (82–144.8 µM), lysis could not be recorded due to its rapid action. That is why low concentration (44 µM) was applied to enable recordings. An equal amounts of KTN6 gp46 in 5 mM HEPES/NaOH (pH 7.4) and *P. aeruginosa* PAO1 Pirnay cells were mixed and fixed on a PBS - 0.2% agar pad for observation under the microscope. KTN6 cause complete lysis of the bacterial cells after 16 min (Fig. 8.9).

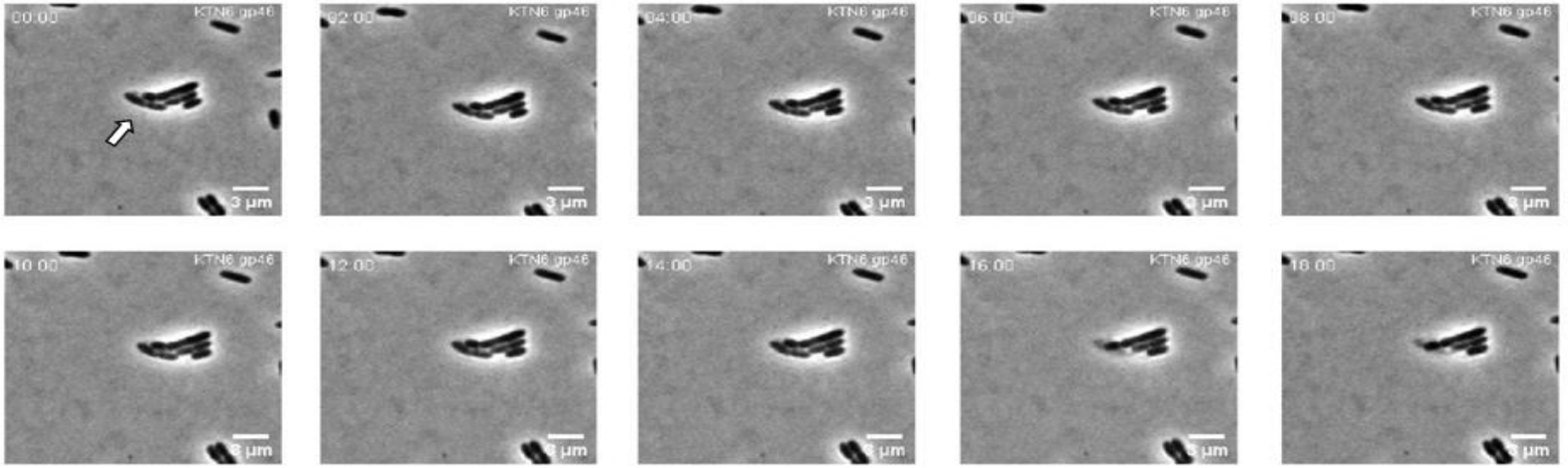


Figure 8.9: Recordings of *P. aeruginosa* Prinay strain lysis caused by KTN6 gp46. White arrow indicates bacterial cell, that was lysed after around 15 min after application of 44 µM of phage enzyme. Scale bare represents 3µm.

8.5. Evaluation of enzyme stability

Some statements and observations were made after recombinant protein preparation:

1) the elution buffer used for purification caused protein precipitation within three days in case of all proteins.
2) 5 mM HEPES pH 7.4 with 30% glycerol prevented precipitation and kept the protein stable at 4°C. Importantly, glycerol is a broadly used stabilizer of proteins at low temperatures since it has cryoprotectant properties.
3) The highest enzymatic activity was achieved, when protein was purified by gravity-flow chromatography at 4°C using a Ni^{2+}-NTA resin. Possibly due to different imidazole concentrations, temperature, as well as higher pressure applied during purification with FPLC.

pH stability of KTN6 gp46

Using OM permeabilized *P. aeruginosa* PAO1 cells suspensions adjusted to a different pH, it was determined that KTN6 gp46 (0.125 μM) is active at pH values ranging from 5 to 9, maintaining between 26.6 ± 10% and 100% of its maximal activity at pH 7 (Fig. 8.10). The highest activity is observed at pH range 6 - 8 with a sharp decrease below pH 6 and above pH 8. At pH 8 the relative activity remains 77.2 ± 3.74% and at pH 6 remains 73.9 ± 25.6%, although for this condition the standard deviation was very high.

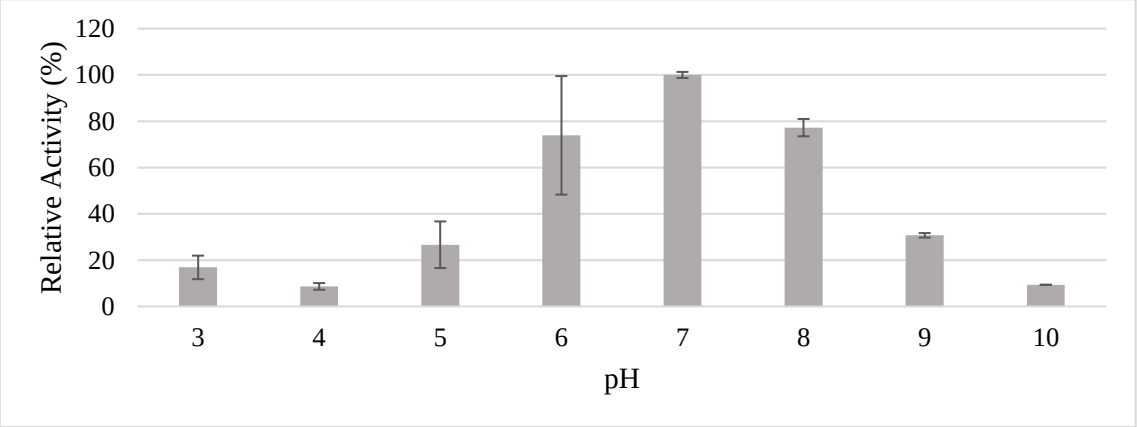


Figure 8.10: The KTN6 gp46 enzymatic activity dependence on pH dependence. Activity is expressed relative to the maximal muralytic activity at the optimal pH (pH 7). Each bar represents the mean of three independent experiments, error bars indicate the standard deviation.

8.6. Discussion

The increasing frequency of multi-drug-resistant pathogens becomes a serious public-health problem, due to limited treatment options (Breidenbach Stein *et al.*, 2011; Poole, 2011). The need for new therapeutics is high. Bacteriophages offer an alternative to current antibacterial treatments, however they present also few disadvantages, such as possible DNA transduction or mutation tendency. In 1991 Gasson introduced the concept of endolysins, phage-borne enzymes with bacteriolytic activity towards Gram-positive microbes. In 2001 Nelson *et al.* (2001) proved for the first time their bactericidal effect against *Streptococcus* using a mouse model of infection in the upper respiratory tract. Since that time many endolysins against Gram-positive pathogens have been studied. In the case of Gram-negative pathogens situation remained for a long time unsolved and neglected (Nelson *et al.*, 2012). Only in recent years, new ideas and strategies to overcome OM barrier were presented, allowing successful administration of endolysin (Briers *et al.* 2007, 2008, 2012, 2014, 2016; Oliveira *et al.*, 2014).

In this chapter, we described six PG hydrolases encoded by *P. aeruginosa* phages. Among which two are VAPGHs associated with phage tail structure (KT28 gp41 and PA5oct gp214) and four are endolysins, expressed in the end of lytic cycle (KT28 gp49, KTN6 gp46, PA5oct gp250 and KTN4 gp48). *In silico* analysis indicated that all tested proteins have globular structure, which is a common feature of PG degrading enzymes isolated from phages infecting Gram-negative bacteria (Briers *et al.*, 2007; Walmagh *et al.*, 2013). The prediction of secondary structure elements and tertiary conformation allowed for prediction of KT28 gp41 domain, which was implemented in analysis due to difficulties in full length protein expression. Moreover, this analysis in combination with a ligand binding sites prediction provided suitable targets for proteins engineering with the use of e.g. side-directed mutagenesis in order to alter protein specificity or influence protein stability, that could be of interest in food technology, where processing at high temperatures occurs (e.g. pasteurization). In the case of PA5oct gp214 and KTN4 gp48 coverage score was below 30%, and achieved structures cannot be trusted. In depth protein structure analysis with the use of e.g. crystallography, would also improve understanding of enzyme function and evaluate the presented analysis.

Protein expression of KT28 gp41 and KTN4 gp48 was difficult with occurrence of inclusion bodies. Possibly introduction of an alternative expression system, such as yeast system, could solve protein toxicity problem. Further, protein purification

required implementation of different chromatography steps. An suitable alternative could be the use of e.g. Dynabeads® coated in a cobalt-based Immobilized Metal Affinity Chromatography (IMAC) chemistry (Life Technologies), which binds histidine-tagged proteins with higher selectivity. Another possibility is an alternation of His-tag to e.g. Strep-tag.

The antibacterial activity of phage recombinant proteins was analyzed by use of several techniques. The PG degradation ability of all enzymes was first confirmed by zymography, that also excluded the potential influence on PG of other co-purified proteins. The muralytic activity with the use of zymograms was also reported in the case of e.g. LysK and endolysin 555 from RcGTA phage (Westbye *et al.*, 2014; Becker *et al.*, 2009; Westbye *et al.*, 2013). According to antimicrobial assay the most potent enzyme proven to be KTN6 gp46. This protein had only 4 amino acid residues difference in compare to KT28 gp49 amino acid sequence, however its bactericidal activity was almost two times stronger. Further investigation is required to establish if amino acid sequence alterations caused changes of protein tertiary structure conformation and enzyme activity or observed differences are result of recombinant protein preparations. Importantly, a similar observation was made by Walmagh *et al.* (2013), who reported different results for highly similar (94%) PsP3 gp10 and P2 gp09 enzymes. In presented study antibacterial activity of all six proteins was in a range of 1.0 ± 0.5 to 3.0 ± 0.3 log reduction of bacterial CFUs. Walmagh *et al.* (2013) and Oliveira *et al.* (2014) presented also similar bactericidal effect of other endolysins derived from Gram-negative infecting phages, where reduction in log-units ranged from a 1.89 to a 3.08. However, modular endolysins showed a greater antibacterial activity against *P. aeruginosa* PAO1 ranging from a 4.36 to a 2.74 log reduction (Walmagh *et al.*, 2013 and Rodriguez-Rubio *et al.*, 2016). This presents a high potential of modular endolysins. With the power of protein engineering such enzymes could be manufactured by fusion of different PG degrading enzymes domains.

If we compare enzymatic activity of KTN6 gp46 with other endolysins analyzed using the same method under optimal conditions, its specific activity (11,129 units/μM) is the second best activity from all of the globular endolysins tested and the fifth best activity if modular endolysins are also taken into account (Table 8.5). However, the assay was conducted in 1:10 dilutions, which gave the rough indication of activity. The slope should be calculated with more narrow dilutions, like a 1:2 dilution, as it was performed by Briers *et al.* (2007) and Rodriguez-Rubio *et al.* (2016). It also has to be stressed, that

antibacterial and enzymatic assays do not correlate with each other. This was also observed by Rodriguez-Rubio *et al*. (2014) in the case of gp110 from *Salmonella* phage 10, that presented the highest specific activity (34,240 units/μM) for OM permeabilized *P. aeruginosa* PAO1, but a moderate antibacterial activity (2.74 ± 0.11) to PAO1. A plausible explanation could be the difference in permeability of the outer membrane after chloroform and EDTA treatment. Chloroform disrupts the outer membrane more severely than EDTA does, making the peptidoglycan more accessible for the endolysin (Lazaroaie *et al.* 2010; Briers *et al.*, 2008).

Table 8.5: List of modular and globular endolysins ranked in decreasing activity (units/μM) (Rodriguez-Rubio *et al.*, 2016).

Endolysin	Activity (units/ μM)	Structure
Gp110	34240	Modular
OBPgp279	19979	Modular
LysEC8	17103	Globular
PVP-SE1gp146	13614	Modular
EL188	4735	Modular
201φ2-1gp229	4469	Modular
KZ144	2058	Modular
PsP3gp10	1380	Globular
P2gp09	829	Globular
BcepC6Bgp22	786	Globular
Lys68	400	Globular
CR8gp3.5	315	Globular
K11gp3.5	134	Globular
KP32gp15	1117	Globular
LysAci7	94	Globular

Further, the host range was tested on *E. coli* and *S.* Enteritids strains, however no effect was observed. Although the peptidoglycan layer of all Gram-negative pathogens is chemically the same (A1γ-chemotype), the permeabilization effect of EDTA on the outer membrane is different. EDTA binds the stabilizing divalent cations in the LPS layer, thereby destabilizing the OM. Due to the lower degree of LPS phosphorylation and thus a lower amount of stabilizing divalent cations of the OM of *E. coli* and *Salmonella* in compare to *P. aeruginosa*, their OM is less prone to the action of EDTA (Schleifer *et al.*, 1972; Loessner, 2005). Similar results were also observed with other previously characterized endolysins (Walmagh *et al.*, 2012; Rodríguez-Rubio *et al.*, 2016; Oliveira *et al.*, 2014; Lim *et al.*, 2014). Oliveira *et al*. (2014) suggested new OM permeabilizers, a citric acid or malic acid, however in case of our PG hydrolases this did not improve the results. Alakomi *et al*. (2003) reported also different response of *S.* Typhimurium to EDTA treatment along the logarithmic growth. Bacteria in early-exponential phase were more sensitive in compare to mid- and late-exponential phase.

Moreover, the increase of cross-linked bonds in the peptidoglycan layer of bacteria in stationary phase can interfere with the antibacterial activity of endolysins.

The optimal pH for KTN6 gp46 is pH 7. This neutral pH for KTN6 gp46 activity is consistent with other previously characterized Gram-negative endolysins (Loessner *et al.*, 2005; Walmagh *et al.*, 2012; Oliveira *et al.*, 2014 and Briers *et al.*, 2007), such as a globular endolysin Lys68 isolated from phage lytic towards *S.* Enteritidis (Oliveira *et al.* 2014). The drop in activity at pH below pH 6 and above pH 8 is also comparable to the observations recently made for endolysin gp110 from *Salmonella* phage 10 (Rodriguez-Rubio *et al*. 2016).

The presented study is only the initial analysis PG degrading enzymes activity, further experiments elucidating e.g. its cytotoxicity, bactericidal effect in combination with other antimicrobials, the resistance occurrence, as well as detailed biochemistry analysis would be essential. Nevertheless, the features of KTN6 gp46 makes it a good candidate for future engineering with antibacterial peptides ('Artilysation'). This approach has the possibility to further improve its lytic activity and expand the host spectrum. It has already been demonstrated that fusion with a CBD improve the antimicrobial activity of PG hydrolases (Rodriguez-Rubio *et al.*. 2012). In Rodríguez-Rubio *et al.* (2012) study, a 4.8 - fold increase of the HydH5 lytic activity was observed when one of the catalytic domains was fused to a lysostaphin SH3b CBD.

Chapter 9
General discussion

"The more you know, the more you know you don't know."

Aristotle

9.1. Bacteriophage research

The discovery of the first antibiotics, Salvarsan, Prontosil and penicillin at the beginning of 20th century gave hope to the world to end the bacterial infectious diseases (Almov, 2010). However, only a few years later the presence of the first bacterial resistance was reported. In years 1950-1970 there was a "golden era" for identification and development of new antibiotics, however inappropriate therapy and drug overuse led to a selection of multidrug resistance bacteria. Currently, antibiotics still save many lives, but there are many hurdles in the new antibiotic development, especially against resistant Gram-negative pathogens. Thus there is an urgent demand for novel and effective strategies to combat bacterial infections. The vast bacteriophage diversity, low cost production and their high efficacy against the target bacteria, make phages and their derived products very attractive alternative/complementary antibacterial therapies.

The therapeutic features of bacteriophages were already noticed at the beginning of 20th century. Unfortunately, due to antibiotic era most of studies and phage therapy were continued only in Georgia, Russia and Poland. Recently "old-new" therapy is again intensively studied, yet most of the research focus on *in vitro* evaluation. There is only a small number of *in vivo* studies, especially focused on phage interaction with immune system, phage safety and bacterial resistance analysis. For this reason bacteriophages currently are not allowed in Europe for therapy. Thus it is important to further explore phage biology, genetics and their safeness.

9.1.1. Microbiological assessment of *P. aeruginosa* phages for therapeutic purposes

The first goal of the present study was to isolate and characterize four *P. aeruginosa* specific phages and evaluate their lytic activity. Phages KT28, KTN6 and KTN4 have a broad host range against several *P. aeruginosa* clinical strains, in contrast to novel giant PA5oct which exhibited a narrow lytic spectrum. Furthermore, testing of their efficacy against biofilm and *Pseudomonas* infection of healthy and CF lung epithelia cell lines revealed a potent phage activity. This feature could be a good indicator for therapeutic purposes. Further tests with human lung epithelia cells would be required,

especially with the use of primary cell cultures, isolated from healthy individuals and CF patients without extensive passages, to retain as much as possible their natural properties. Additionally, the ability of the lung mucus penetration by bacteria and phage should be tackled it detail, by introducing e.g. fluorescent tag, that could be visualized using microscopy. Furthermore, the experimental animal lung infection model could be applied to study influence of phage therapy *in vivo.*

The biggest hurdle remain the biofilm degradation experiments, due to lack of highly sensitive and reproducible techniques with low standard deviations. The Drulis-Kawa group and collaborators address this problem by applying new alternative assays including interferometry, goniometry and profilometry (Danis-Wlodarczyk *et al.*, 2015 and 2016). Biofilm research should be also further extended by analysis of phage cocktails efficacy and study of mixed biofilms, that would mimic more natural multispecies biofilms.

9.1.2. Phage genome analysis

The functional annotation of phage genomes is a crucial step to assess phage safety and to reveal and explain basic phage biology properties. In our study, we focused on four phage genomes. A first set of ORFs was identified and gene features were assigned based on genome sequence homology. No unwanted genetic elements (e.g. associated with lysogeny or virulence) were identified. However, for more than half of predicted genes possible function could not be assigned. These proteins could have an impact on phage biology and different aspects of bacterial physiology. Unraveling biological functions of these unknown proteins is a major importance and it could be elucidated by e.g. screening the libraries of unknown genes toward e.g. inhibitory, antibiofilm or metabolism modulation functions. Unfortunately, this research is complex and time consuming. Nevertheless, it could increase availability of new phage-born therapeutics and allow to design the high-throughput screens with targeted host process to improve the selection of desirable phage proteins (Van den Bossche *et al.*, 2014; Wagemans *et al.*, 2014).

Moreover, elucidation of protein structure by NMR, crystallography or CryoEM analysis, would provide important new information about the interaction site and allow protein docking studies. Recently more studies address this problem, but still detailed analysis are required and a systematic approach is lacking (Van den Bossche *et al.*, 2014; De Smet *et al.*, 2016).

Suitable tools to address the identification of genes and gene products include ESI-MS/MS and RNA seq, which drastically impact genome annotations by providing experimental confirmation. Mass spectrometry provides an identification of individual proteins, but no functional insights can be gathered. RNA seq was performed on the novel giant phage PA5oct, which confirmed the presence and expression of a vast number of 'ORFan' genes (without sequence homology). This technique can also lead to the elucidation of regulatory elements and gene expression pattern, as well as evaluate the impact of phage infection on transcription regulation of host genes. Particularly early gene features could be assigned, involved in the shutdown of host metabolism through various mechanisms, e.g. redirection of host RNA polymerase to phage middle and late promoters (Calendar, 2006; Rothman-Denes, 2015; Chevallereau *et al.*, 2016). The *P. aeruginosa* transcription apparatus was rapidly overwhelmed with viral protein transcription. This has been previously reported also in case of e.g. phage T4 and PAK_P3 (Ueno *et al.*, 2004; Uzan, 2009; Chevallereau *et al.*, 2016). The big challenge of RNA seq remains the cost as well as the manual curation needed to process the data. The correct storage and visualization of this information also remains a challenge.

9.1.3. Phage evolution

Comparative genomics allowed for the implementation of a genome-based mathematical model for reticulate phylogenomic networks based on ortholog clustering (Lima-Mendez *et al.*, 2008; Leplae *et al.*, 2010; Jang *et al.*, 2013). Subsequent elucidation of a common ancestors of phage families, phage evolution and diversity at the genomic and amino acid sequence level is possible. The pool of shared proteins/genes was reconstructed from the pattern of protein families included in a pool of homologous proteins identified in different phages (Jang *et al.*, 2013). In presented dissertation, phages KT28 and KTN6 had the highest similarity to *Pbunavirus*, particularly with *Pseudomonas* phage LMA2, whereas KTN4 forms an in-group relationship with φKZ, φPA3, and 201φ2-1 phages. However, novel giant phage PA5oct showed only small pool of protein homology to *T4virus* and proved that the diversity of phages is not yet „saturated". The question is how many phage genera remain to be discovered? Well, with global amount of phages estimated at 10^{31} particles any estimation is purely speculative, although closely related phage can be found on distinct geographical locations (Savalia *et al.*, 2008). However, the standard phage identification and propagation procedures appears be insufficient for isolation and identification of all bacteriophage species. Chen

et al. (2001) estimated, that less than 1% of the phage species residing in river samples could be detected by standard plaque assay. The viral metagenomics present a new approach of phage diversity assessment. This method allows for sequencing of total viral components present in a particular environment sample. Moreover, the most dominant viral particles can be elucidated in particular location, without phage propagation step as well as presence of phages, that are not amenable for propagation or simply do not have a host in culture (Clokie *et al.*, 2011).

In my opinion, the biggest hurdle of genome analysis today is the lack of general rules or guidelines for annotation, especially for the gene feature naming, that leads to accumulation of homologues genes with different names. Unfortunately, efforts of e.g. ICTV (Kropinski *et al.*, 2009) to systematize annotation, classification and phage naming are sometimes ignored, which complicates data information available in databases. Moreover, number of NCBI submitted genomes are sometimes annotated inaccurately, which is misleading and masks promising data.

9.1.4. The future of bacteriophage therapy

Phages exhibit several advantages for therapeutic purposes, such as high diversity and abundance, and bacteriolytic activity. In addition, they show a high specificity to one genus or species, synergy with other phages with different lysis mechanisms and/or bacterial surface receptors in cocktail preparations, as well as with antibiotics. Moreover, phage therapy is self-dosing and phage preparations are not expensive (Drulis-Kawa *et al.*, 2012 and 2015; Hyman *et al.*, 2012; Bikard *et al.*, 2014; Brown-Jaque *et al.*, 2015; Roach *et al*, 2015). To date, bacteriophages have proven to be useful in the food processing and show encouraging results in animal models. Also, many years of phage therapy conducted on humans in Tbilisi (Georgia) and Wrocław (Poland) presented high success rate of this alternative therapy. In Europe, Pherecydes Pharma launched randomized controlled single-blind clinical study in I/II trial phase of two phage cocktails tested on serious burn victims in 2013. This is the first international clinical phage therapy study, that introduce international standards of clinical assessment (www.pherecydes-pharma.com). These efforts indicate that it is possible to move phage therapy forward under the current legislative requirements. However, it is of a crucial importance for the competent public authorities to formulate a tailored legislative framework that encourages companies to invest in bacteriophage therapy and formulate appropriate solutions in a classical pharmaceutical context (clinical trials, production requirements

etc.). Hopefully, a rational bacteriophage therapy concept will be developed, to avoid historical mistakes that occurred in case of antibiotic therapy (Kegan, 1988, Miller, 1996).

9.2. Harvesting of antibacterial enzymes

Phage therapy also has distinct disadvantages, including the vast pool of unknown and potentially unwanted proteins, such as toxins or integrases, that could lead to lysogeny state and horizontal gene transfer between phages and its host (e.g. transfer of virulence genes) via specialized or generalized transduction. Moreover, bacterial resistance to phages was also noticed previously (Labrie *et al*., 2010; Danis-Wlodarczyk *et al.*, 2015). The Drulis-Kawa group in their biofilm eradication assays observed the rapid appearance of phage-resistant variants (Danis-Wlodarczyk *et al.*, 2015). We have presented our preliminary results during the 21st Biennial Evergreen International Phage Meeting (2015) (Drulis-Kawa *et al.*, unpublished results), regarding the emergence of phage-resistant *P. aeruginosa* PAO1 mutants obtained from biofilms treated with KT28 and KTN6 (LPS-dependent) phages, as well as for KTN4 and φKZ (pili-dependent) phages, which exhibited the lost or truncation in LPS O-antigen structure and reduction in twitching motility, respectively. The receptor modification is the first line defense mechanism of *P. aeruginosa* in response to phage presence (Labrie *et al.*, 2010). Moreover, bacterial resistance occurred faster in bacteria embedded in biofilms, compared to their planktonic counterparts.

To circumvent phage therapy hurdles, phage-borne enzymes with antibacterial properties could be introduced as antimicrobials. This seems to be an alternative solution in compare to phage formulations. In this study we presented two types of proteins, PG degrading enzymes and polysaccharide depolymerases.

9.2.1. Peptidoglycan degrading enzymes

PG degrading enzymes, including virion-associated PG hydrolases and endolysins, have muralytic activity and are responsible for host cell wall degradation during infection and phage progeny release at the end of infection cycle, respectively. Several studies have already proven their antibacterial potential against Gram-positive and Gram-negative bacteria *in vitro* and *in vivo.* Despite conserved biological function, PG degrading enzymes are heterogeneous in structure. This divergence is directly related to the Gram-positive and Gram-negative bacteria cell wall structure. They can be globular

or modular enzymes. The first type, mainly correlated with Gram-negative infecting phages, consist a single EAD, that cleaves specific bonds within PG (Briers *et al.*, 2007). Second type, often related to Gram-positive infecting phages, have an N-terminal EAD and C-terminal cell wall-binding domain (CBD) (Fischetti, 2010; Nelson *et al.*, 2012) with few exceptions.

The antibacterial tests with *P. aeruginosa* and equal concentrations of both types of endolysins revealed that modular structures can have up to five-fold greater activity in compare to their globular counterparts (Walmagh *et al.*, 2013). Also fusion of the CBD to a virion-associated endolysin can triple the activity (Briers *et al.*, 2015). The superior efficiency of modular endolysins with a CBD can be explained by its ability to cleave and increase insoluble substrate proximity, positioning it into enzyme active site, as well as improvement of enzyme kinetics (Alberghina, 2005).

Gram-negative bacteria have an OM which is a highly effective permeability barrier. To overcome this obstacle, PG degrading enzymes must be delivered in combination with a high hydrostatic pressure (150 - 200 MPa) or OM permeabilizers, such as EDTA, citric acid or malic acid (Briers *et al.*, 2011; Oliveira *et al.*, 2014), which not always work sufficiently. Briers *et al.* (2014a) presented another approach focusing on engineered endolysins named 'Artilysin®s', that consists of a recombinant modular endolysin and an antimicrobial peptide, enabling transport through the OM. This peptide can either be polycationic, to disrupt the stabilizing ionic bonds between divalent cations and the negatively charged phosphate groups, or hydrophobic, to interfere with the hydrophobic moieties present in the lipid A part of LPS (Briers *et al.*, 2014). Moreover, recombinant PG degrading enzymes isolated from phages targeting Gram-negative bacteria as globular proteins have an theoretically similar enzymatic activity towards the peptidoglycan of most Gram-negative bacteria, due to a highly conserved A1γ peptidoglycan structure (Schleifer *et al.*, 1972). In contrary, the proteins derived from Gram-positive targeting phages, are more bacterial strain/species specific, due to the presence of cell wall binding domains presented in most of those enzymes, which condition a highly specific recognition of particular targets in the cell wall.

Importantly, to date there are no reports describing bacterial resistance mechanisms or lost sensitivity to 'lysis from without' (Roach *et al.*, 2015). Several studies failed in resistance generation *in vitro* by repeated pathogen exposure to sub-lethal concentrations of endolysins or endolysin-sensitive strains mutagenesis (Schuch *et al.*, 2002). No resistance occurred or resistant mutants were not promoted, while in control

groups, treated with only antibiotics, resistance increased 10 - 100,000 fold. Also several studies on animal models confirmed this *in vitro* observations (e.g. Nelson *et al.*, 2001; Schuch *et al.*, 2002; Cheng *et al.*, 2005; Yiin *et al.*, 2006). For example no resistance occurred after treatment of MDR *S. aureus* by MV-L lysin derived from phage φMR11 with the use of mouse infection model and pathogen was efficiently eradicated (Rashel *et al.*, 2007).

Identification of genes encoding these proteins in phage genomes is not complicated due to the relatively strong conserved protein domains. Furthermore, in many cases of PG degrading enzymes of Gram-negative origin, protein expression and purification is relatively straightforward, with the exception of Signal-Arrest-Release (SAR) endolysins, where protein expression is more complicated (Briers *et al.*, 2011). The PG degrading phage enzymes act very rapidly with the microbes being lysed in few seconds after exogenous application of even a small amount of purified endolysins (Briers *et al.*, 2014).

In our study we analyzed six enzymes, including four endolysins (KT28 gp49, KTN6 gp46, PA5oct gp250, KTN4 gp48) and two VAPGHs (KT28 gp41, PA5oct gp214). All presented globular structure. The most active was endolysin KTN6 gp46, that showed high efficiency in combination with EDTA. Further efforts should be focused on structural studies to gain insight into e.g. substrate binding site, thermostability, that could be improved with use of side-directed mutagenesis. Also, as mentioned before, protein could be fused with antimicrobial peptide to avoid EDTA/enzyme mixture. Moreover, it would be interesting to test susceptibility of the range of Gram-negative pathogens to engineered enzyme with or without EDTA.

Predatory PG degrading enzymes are currently of high interest as alternative microbials (Briers *et al.*, 2014). They could be use e.g. externally on wounds in the form of ointments or supplementary compounds in bandages or other types of dressings as well as biocontrol of animal/plants diseases and food preparation (Kim *et al.*, 2014; Schmelcher *et al.*, 2014). There is still little know about their immunogenicity, which potentially could limit their internal application. The CBDs are also consider for rapid detection of pathogens to prevent food-borne outbreaks. Importantly, with the capabilities of modern synthetic biology and molecular engineering, these enzymes could be designed and adjusted to specific purposes. The ability to produce them as recombinant proteins allows to optimize manufacturing and purification processes, that could be adapted from currently existing methods (e.g. those used to produce insulin).

9.2.2. EPS degrading enzymes

P. aeruginosa produces EPS in the form of capsules or a thick biofilm layer (Campodonico *et al.*, 2008). EPS serve as molecular patterns for recognition by the innate immune system, and also provide shields to antibiotic entry and host defence mechanisms (e.g. complement system and phagocytosis). Specific bacteriophages contain EPS depolymerases as virion-associated proteins which serve to degrade EPS to enable infection. Such phages can be identified by the appearance of a halo zone around the phage plaques, which results from the enzymatic degradation of bacterial EPS without phage infection (Azeredo *et al.*, 2008). However, the knowledge of EPS depolymerases is currently very limited and mainly derived from *E. coli* phages (e.g., Leiman *et al.*, 2007). Not much is known of their complicate structure, including a substrate binding or cleaving domains. The PDB database contains only few structures of EPS depolymerases. It is also unknown to what extent the depolymerase activity is required for the infection process and how possible variations in the amount and density of the polysaccharide influence this process (Pelkonen *et al.*, 1992). There is also lack of biochemical studies, which would revealed protein mode of action and specificity.

In our research we presented the process of EPS depolymerase identification and recombinant protein preparation. The first obstacle was the limited protein sequence similarity of these enzymes in the public databases and their limited, general or even incorrect annotation. Further, these proteins are relatively large, ranging to >1000 amino acids, hampering efficient cloning, expression and proper folding. To achieve a successful expression, different expression systems and conditions needed to be evaluated and further analyses beyond what's been evaluated here can be assessed.

In our study we decided to focus only on enzyme domains, that had shorter amino acid sequence and their handling was much easier. The decision how to define those domains was not easy, due to lack of EPS depolymerase structural studies. The efforts by Cornelissen and Leiman to elucidate LKA1 gp49 structure allowed us to more precisely define position of a LKA1 gp49 domain. The stability tests suggested that, possibly the C-terminal part of LKA1 gp49 domain should be extended, that proved to be responsible for protein folding and trimerization processes, as was described elsewhere (Schwarzer *et al*., 2007; Walter *et al*., 2008; Xiang *et al*., 2009).

The activity range of these phage enzymes should be studied in details against different isolates and species. Cornelissen *et al.* (2011) presented that an EPS depolymerase has a high specificity, limited to the EPS produced by the specific bacterial strain used for phage propagation.

Future perspectives

The applications of phage-borne EPS degrading enzymes is an innovative approach that exploits the weapons of the bacteriophage. Such enzymes could offer a novel stand-alone strategy to control biofilm growth, which will be also complementary to existing strategies. Unfortunately, extensive fundamental and technology-development research remains to be done to allow a proper assessment of the potential of these enzymes, including the biochemical analysis of the EPS diversity and the development of sensitive, quantitative and reproducible biofilm assays.

Furthermore, the narrow spectrum of EPS depolymerases could be overcome by protein subjection to *in vitro* direct evolution, in order to tailor them as optimized agents that prevent biofilm formation or/and eradicate existing biofilms. Directed evolution can broaden the substrate specificity of the enzymes and proteins can be tailored to degrade the greatest possible number of biofilm of clinical *P. aeruginosa* isolates, including those of cystic fibrosis infection. The power of *in vitro* evolution can be used to 'relax' a natural specificity and to create broad-spectrum EPS depolymerases.

References

Abbasifar R, Griffiths MW, Sabour PM, Ackermann HW, Vandersteegen K, Lavigne R, Noben JP, (2014) Alanis Villa A, Abbasifar A, Nash JH, Kropinski AM., (2014), Supersize me: Cronobacter sakazakii phage GAP32. Virology.;460-461:138-46.

Abbott DW, Boraston AB, (2007) The structural basis for exopolygalacturonase activity in a family 28 glycoside hydrolase. J Mol Biol.;368(5):1215-22.

Abedon ST, (2011) Bacteriophages and biofilms: ecology, phage therapy, plaques. Nova Science Publishers, Inc. New York, 1-50.

Abeyrathne PD, Daniels C, Poon KK, Matewish MJ, Lam JS, (2005) Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. J Bacteriol.;187(9):3002-12.

Ackermann HW, (2003) Bacteriophage observations and evolution. Res. Microbiol. 154, 245–251.

Ackermann HW, (2007) 5500 Phages examined in the electron microscope. Arch Virol.;152(2):227-43.

Ackermann HW, Auclair P, Basavarajappa S, Emadi Konjin HP, Savanurmath C, (1994) Bacteriophages from *Bombyx mori*. Arch Virol.;137:185–190.

Ackermann HW, Cartier C, Slopek S, Vieu JF, (1988) Morphology of *Pseudomonas aeruginosa* typing phages of the Lindberg set. Ann Inst Pasteur Virol.;139(4):389-404.

Adams MH, (1959) Bacteriophages. In: Adams MH, editor. Bacteriophages. New York: Interscience Publishers; 27–30.

Adriaenssens EM, Mattheus W, Cornelissen A, Shaburova O, Krylov VN, Kropinski AM, Lavigne R, (2012) Complete genome sequence of the giant *Pseudomonas* phage Lu11. J Virol.;86(11):6369-70.

Agol, VI, (1974) Towards the system of viruses. Biosystems 6(2),113–132.

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF, (2004) GDSL family of serine esterases/lipases. Prog Lipid Res.;43(6):534-52.

Alakomi HL, Saarela M, Helander IM, (2003) Effect of EDTA on *Salmonella enterica* serovar Typhimurium involves a component not assignable to lipopolysaccharide release. Microbiology, 49(Pt 8):2015-21.

Alberghina L, (2005) Protein engineering in industrial biotechnology. Milano: Harwood Academic.

Alcorn JF, Wright JR, (2004) Degradation of pulmonary surfactant protein D by *Pseudomonas aeruginosa* elastase abrogates innateimmune function. J Biol Chem.;279(29):30871-9.

Allen L, Dockrell DH, Pattery T, Lee DG, Cornelis P, Hellewell PG, Whyte MK, (2005) Pyocyanin production by *Pseudomonas aeruginosa* induces neutrophil apoptosis and impairs neutrophil-mediated host defenses *in vivo*. J Immunol.;174: 3643–3649.

Altmann F, Christian R, Czerny T, Nimmich W, Marz L, (1990) Bacteriophageassociated glycan hydrolases specific for *Escherichia coli* capsular serotype K12. Eur. J. Biochem. 185, 307–312.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, (1990) Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Amiel E, Lovewell RR, O'Toole GA, Hogan DA, Berwin B, (2010) *Pseudomonas aeruginosa* evasion of phagocytosis is mediated by loss of swimming motility and is independent of flagellum expression. Infect. Immun. 78:2937–2945.

Aminov RI, (2010) A brief history of the antibiotic era: lessons learned and challenges for the future. Front Microbiol.;1:134.

Anderson DM, Frank DW, (2012) Five mechanisms of manipulation by bacterial effectors: a ubiquitous theme. PLoS Pathog. 8:e1002823.

Aryal S, (2015) Differences between Gram Positive and Gram Negative Bacteria [Online]. http://www.microbiologyinfo.com/differences-between-gram-positive-and-gramnegative-bacteria

Assenov Y, Ramírez F, Schelhorn SESE, Lengauer T, Albrecht M, (2008) Computing topological parameters of biological networks. Bioinformatics.;24: 282–284.

Augustin M, Ali-Vehmas T, Atroshi F, (2004) Assessment of enzymatic cleaning agents and disinfectants against bacterial biofilms. J. Pharm. Pharm. Sci., 7, 55-64.

Aylett CH, Izoré T, Amos LA, Löwe J, (2013) Structure of the tubulin/FtsZ-like protein TubZ from *Pseudomonas* bacteriophage ΦKZ. J Mol Biol.;425(12):2164-73.

Azeredo J, Sutherland IW, (2008) The use of phages to the removal infectious biofilms. Curr. Pharm. Biotech.,9, 261-266.

Bailey TL, Gribskov M, (1998) Combining evidence using p-values: application to sequence homology searches. Bioinformatics.; 14: 48–54.

Balasubramanian D, Schneper L, Kumari H, Mathee K, (2013) A dynamic and intricate regulatory network determines *Pseudomonas aeruginosa* virulence. Nucl. Acids Res. 41, 1–20.

Ballok AE, O'Toole GA, (2013) Pouring salt on a wound: *Pseudomonas aeruginosa* virulence factors alter Na+ and Cl- flux in the lung. J Bacteriol. ;195(18):4013-9.

Baltimore D, (1971) Expressionofanimalvirusgenomes.Bacteriol.Rev.35(3), 235–241.

Bansal S, Harjai K, Chhibber S, (2014) Depolymerase improves gentamicin efficacy during *Klebsiella pneumoniae* induced murine infection. BMC Infect. Dis., 14, pp. 1–11.

Barbieri JT, Sun J, (2004) *Pseudomonas aeruginosa* ExoS and ExoT. Rev. Physiol. Biochem. Pharmacol. 152:79 –92.

Barbirz S, Müller JJ, Uetrecht C, Clark AJ, Heinemann U, Seckler R, (2008) Crystal structure of *Eschericha coli* phage HK620 tail spike: podoviral tailspike endoglycosidase modules are evolutionary related; Mol Microbiol.;69(2):303-16.

Barker AP, Vasil AI, Filloux A, Ball G, Wilderman PJ, Vasil ML, (2004) A novel extracellular phospholipase C of *Pseudomonas aeruginosa* is required for phospholipid chemotaxis. Mol. Microbiol. 53:1089 –1098.

Battin TJ, Sloan WT, Kjelleberg S, Daims H, Head IM, Curtis TP, Eberl L, (2007) Microbial landscapes: new paths to biofilm research. Nat Rev Microbiol.;5(1):76-81.

Becker SC, Dong S, Baker JR, Foster-Frey J, Pritchard DG, Donovan DM, (2009) LysK CHAP endopeptidase domain is required for lysis of live staphylococcal cells. FEMS Microbiol. Lett., 294(1):52-60.

Bernad A, Blanco L, Lázaro JM, Martín G, Salas M, (1989) A conserved 3'----5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell.;59(1):219-28.

Besemer J, Lomsadze A, Borodovsky M, (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29, 2607–2618.

Bieńkowska-Szewczyk K, Taylor A, (1980) Murein transglycosylase from phage lambda lysate. Purification and properties. Biochim Biophys Acta.;615(2):489-96.

Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, Duportet X, Fischetti VA, Marraffini LA, (2014) Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. Nat Biotechnol; 32:1146-50.

Blair JM, Webber MA, Baylay AJ, Ogbolu DO, Piddock LJ (2015) Molecular mechanisms of antibiotic resistance. Nat Rev Microbiol.;13(1):42-51.

Bleves S, Viarre V, Salacha R, Michel GP, Filloux A, Voulhoux R, (2010) Protein secretion systems in *Pseudomonas aeruginosa*: a wealth of pathogenic weapons. Int. J. Med. Microbiol. 300:534 –543.

Bonvillain RW, Painter RG, Adams DE, Viswanathan A, Lanson NA, Jr, Wang G, (2010) RNA interference against CFTR affects HL60-derived neutrophil microbicidal function. Free Radic. Biol. Med. 49:1872–1880.

Born Y, Fieseler L, Klumpp J, Eugster MR, Zurfluh K, Duffy B, Loessner MJ. (2014) The tail-associated depolymerase of *Erwinia amylovora* phage L1 mediates host cell adsorption and enzymatic capsule removal, which can enhance infection by other phage. Environ Microbiol.;16(7):2168-80.

Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B, Bartlett J, (2009) Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. Clin Infect Dis.;48(1):1-12.

Bragonzi A, Paroni M, Nonis A, Cramer N, Montanari S, Rejman J, Di Serio C, Döring G, Tümmler B, (2009), *Pseudomonas aeruginosa* microevolution during cystic fibrosis lung infection establishes clones with adapted virulence. Am J Respir Crit Care Med.;180: 138–145.

Breidenstein EB, de la Fuente-Núñez C, Hancock RE, (2011) *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends in Microbiology* 19(8): 419-426.

Brick DJ, Brumlik MJ, Buckley JT, Cao JX, Davies PC, Misra S, Tranbarger TJ, Upton C, (1995) A new family of lipolytic plant enzymes with members in rice, Arabidopsis and maize. FEBS Lett 377:475-480.

Briers Y, Cornelissen A, Aertsen A, Hertveldt K, Michiels CW, Volckaert G, Lavigne R, (2008) Analysis of outer membrane permeability of *Pseudomonas aeruginosa* and bactericidal activity of endolysinsKZ144 and EL188 under high hydrostatic pressure. FEMS Microbiol Lett.;280(1):113-9.

Briers Y, Lavigne R, (2015) Breaking barriers: expansion of the use of endolysins as novel antibacterials against Gram-negative bacteria. Future Microbiol.;10(3):377-90.

Briers Y, Lavigne R, Volckaert G, Hertveldt K, (2007) A standardized approach for accurate quantification of murein hydrolase activity in high-throughput assays. J Biochem Biophys Methods.;70(3):531-3.

Briers Y, Miroshnikov K, Chertkov O, Nekrasov A, Mesyanzhinov V, Volckaert G, Lavigne R, (2008) The structural peptidoglycan hydrolase gp181 of bacteriophage φKZ. Biochem Biophys Res Commun.; 374:747–751.

Briers Y, Peeters LM, Volckaert G, Lavigne R, (2011) The lysis cassette of bacteriophage ϕKMV encodes a signal-arrest-release endolysin and a pinholin. Bacteriophage.;1(1):25-30.

Briers Y, Schmelcher M, Loessner MJ, Hendrix J, Engelborghs Y, Volckaert G, Lavigne R, (2009), The high-affinity peptidoglycan binding domain of *Pseudomonas* phage endolysin KZ144. Biochem Biophys Res Commun.;383(2):187-91.

Briers Y, Volckaert G, Cornelissen A, Lagaert S, Michiels CW, Hertveldt K, Lavigne R, (2007). Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages φKZ and EL. Mol. Microbiol.;65(5):1334-44.

Briers Y, Walmagh M, Grymonprez B, Biebl M, Pirnay JP, Defraine V, Michiels J, Cenens W, Aertsen A, Miller S, Lavigne R, (2014b) Art-175 is a highly efficient antibacterial against multidrug-resistant strains and persisters of *Pseudomonas aeruginosa*. Antimicrob Agents Chemother.;58(7):3774-84.

Briers Y, Walmagh M, Lavigne R, (2011) Use of bacteriophage endolysin EL188 and outer membrane permeabilizers against *Pseudomonas aeruginosa*. J. Appl. Microbiol. 110:778 –785.

Briers Y, Walmagh M, Van Puyenbroeck V, Cornelissen A, Cenens W, Aertsen A, Oliveira H, Azeredo J, Verween G, Pirnay JP, Miller S,Volckaert G, Lavigne R, (2014a) Engineered endolysin-based "Artilysins" to combat multidrug-resistant gram-negative pathogens. MBio.;5(4):e01379-14.

Brohee S, Faust K, Lima-Mendez G, Vanderstocken G, van Helden J, (2008) Network Analysis Tools: from biological networks to clusters and pathways. Nat. Protoc. 3, 1616–1629.

Brown RW, (1956) Composition of Scientific Words. Smithsonian Institutional Press.

Browning C, Shneider MM, Bowman VD, Schwarzer D, Leiman PG, (2012). Phage pierces the host cell membrane with the iron-loaded spike. Structure 20, 326–339.

Brown-Jaque M, Calero-Caceres W, Muniesa M, (2015) Transfer of antibiotic-resistance genes via phage related mobile elements. Plasmid; 79:1-7; PMID:25597519.

Burns JL, Jonas M, Chi EY, Clark DK, Berger A, Griffith A, (1996) Invasion of respiratory epithelial cells by *Burkholderia* (*Pseudomonas*) *cepacia*. Infect. Immun. 64: 4054–4059.

Butterworth MB, Zhang L, Heidrich EM, Myerburg MM, Thibodeau PH, (2012) Activation of the epithelial sodium channel (ENaC) by the alkaline protease from *Pseudomonas aeruginosa*. J. Biol. Chem. 287: 32556 –32565.

Byfield FJ, Kowalski M, Cruz K, Leszczyńska K, Namiot A, Savage PB, Bucki R, Janmey PA, (2011) Cathelicidin LL-37 increases lung epithelial cell stiffness, decreases transepithelial permeability, and prevents epithelial invasion by *Pseudomonas aeruginosa*. J Immunol.;187(12):6402-9.

Calendar R, (2006) The bacteriophages Oxford; New York: Oxford University Press. xiii, 746.

Campodonico VL, Gadjeva M, Paradis-Bleau C, Uluer A, Pier GB, (2008) Airway epithelial control of *Pseudomonas aeruginosa* infection in cystic fibrosis. Trends Mol Med 14: 120-133.

Carter CD, Parks A, Abuladze T, Li M, Woolston J, Magnone J, Senecal A, Kropinski AM, Sulakvelidze A, (2012) Bacteriophage cocktail significantly reduces *Escherichia coli* O157: H7 contamination of lettuce and beef, but does not protect against recontamination. Bacteriophage.;2: 178–185.

Casjens S, (2003) Prophages and bacterial genomics: What have we learned so far? Mol. Microbiol. 49, 277–300.

Casjens SR, Thuman-Commike PA, (2011) Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology; 411:393-415; PMID:21310457.

Catalão MJ, Gil F, Moniz-Pereira J, São-José C, Pimentel M, (2013) Diversity in bacterial lysis systems: bacteriophages show the way. FEMS Microbiol Rev.;37(4):554-71.

Cenens W, Mebrhatu MT, Makumi A, Ceyssens PJ, Lavigne R, Van Houdt R, Taddei F, Aertsen A, (2013). Expression of a novel P22 ORFan gene reveals the phage carrier state in *Salmonella* Typhimurium. PLoS Genet. 9:e1003269.

Ceyssens PJ, Lavigne R, Mattheus W, Chibeu A, Hertveldt K, Mast J, Robben J, Volckaert G, (2006) Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the KMV subgroup within the T7 supergroup. J. Bacteriol. 188:6924 –6931.

Ceyssens PJ, Mesyanzhinov V, Sykilinda N, Briers Y, Roucourt B, Lavigne R, Robben J, Domashin A, Miroshnikov K, Volckaert G, Hertveldt K, (2008)The genome and structural proteome of YuA, a new *Pseudomonas aeruginosa* phage resembling M6. J Bacteriol.; 190: 1429–1435.

Ceyssens PJ, Minakhin L, Van den Bossche A, Yakunina M, Klimuk E, Blasdel B, De Smet J, Noben JP, Bläsi U, Severinov K, Lavigne R, (2014) Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. J Virol.;88(18):10501-10.

Ceyssens PJ, Miroshnikov K, Mattheus W, Krylov V, Robben J, Noben JP, Vanderschraeghe S, Sykilinda N, Kropinski AM, Volckaert G, Mesyanzhinov V, Lavigne R, (2009) Comparative analysis of the widespread and conserved PB1-like viruses infecting *Pseudomonas aeruginosa*. Environ Microbiol.; 11: 2874–2883.

Chaignon P, Sadovskaya I, Ragunah C, Ramasubbu N, Kaplan JB, Jabbouri S, (2007) Susceptibility of staphylococcal biofilms to enzymatic treatments depends on their chemical composition. Appl. Microbiol. Biotechnol. 2007, 75, 125-132.

Chambers LA, Rollins BM, Tarran R, (2007) Liquid movement across the surface epithelium of large airways. Respir. Physiol. Neurobiol. 159:256 – 270.

Chang HC, Chen CR., Lin JW, Shen GH, Chang KM, Tseng YH, Weng S-F, (2005). Isolation and characterization of novel giant *Stenotrophomonas maltophila* phage ϕSMA5. Appl. Environ. Microbiol. 71, 1387–1393.

Chanishvili N, (2012) Phage therapy–history from Twort and d'Herelle through Soviet experience to current approaches. Adv Virus Res; 83:3-40; PMID:22748807.

Chen F, Lu J, (2002) Genomic sequence and svolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. Appl Environ Microbiol.;68(5):2589-94.

Cheng Q, Nelson D, Zhu S, Fischetti VA, (2005) Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. Antimicrob. Agents Chemother., 49(1), pp. 111–117.

Cheng X, Zhang X, Pflugrath JW, Studier FW, (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. Proc Natl Acad Sci U S A.;91(9):4034-8.

Chevallereau A, Blasdel BG, De Smet J, Monot M, Zimmermann M, Kogadeeva M, Sauer U, Jorth P, Whiteley M, Debarbieux L, Lavigne R, (2016) Next-generation "-omics" approaches reveal a massive alteration of host RNA metabolism during bacteriophage infection of *Pseudomonas aeruginosa*. PLoS Genet.;12(7):e1006134.

Chhibber S, Bansal S, Kaur S, (2015) Disrupting the mixed-species biofilm of *Klebsiella pneumoniae* B5055 and *Pseudomonas aeruginosa* PAO using bacteriophages alone or in combination with xylitol. Microbiology, 161(7):1369-77.

Chmielewski RAN, Frank JF, (2003) Biofilm formation and control in food processing facilities, comprehensive reviews in food science and food safety, 2(1), 22-32.

Chou PY Fasman GD, (1974) Prediction of protein conformation. Biochemistry. 13 (2): 222–245.

Chow MS, Rouf MA, (1983) Isolation and partial characterization of two *Aeromonas hydrophila* bacteriophages. Appl Environ Microbiol.;45(5):1670-6.

Ciofu O, Riis B, Pressler T, Poulsen HE, Hoiby N, (2005) Occurrence of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis patients is associated with the oxidative stress caused by chronic lung inflammation. Antimicrob. Agents Chemother., 49, 2276–2282.

Claverie J M ,Abergel C, (2009) Mimivirus and its virophage.Annu.Rev.Genet.43, 49–66.

Clokie MR, Millard AD, Letarov AV, Heaphy S, (2011) Phages in nature. Bacteriophage.;1(1):31-45.

Coggan KA, Wolfgang MC, (2012) Global regulatory pathways and crosstalk control *Pseudomonas aeruginosa* environmental lifestyle and virulence phenotype. Curr. Issues Mol. Biol. 14, 47–70.

Cohen TS, Prince A, (2012) Cystic fibrosis: a mucosal immunodeficiency syndrome. Nat Med. 5;18(4):509-19.

Cohn F, (1872) Untersuchungen fiber bacterien, Beitrage zur Biologie der Pflanzen 1, Heft 2, 127-224.

Colvin KM, Gordon VD, Murakami K, Borlee BR, Wozniak DJ, Wong GCL, Parsek MR, (2011) The pel polysaccharide can serve a structural and protective role in the biofilm matrix of *Pseudomonas aeruginosa.* PLoS Pathog., 7(1):e1001264.

Cooper CJ, Khan Mirzaei M, Nilsson S, (2016) Adapting drug approval pathways for bacteriophage-based therapeutics. Front Microbiol.;7:1209.

Cornelis P, (2008), *Pseudomonas:* Genomics and Molecular Biology, Caister Academic Press.

Cornelis P, Dingemans J, (2013) *Pseudomonas aeruginosa* adapts its iron uptake strategies in function of the type of infections. Front Cell Infect Microbiol.;3:75.

Cornelis P, Matthijs S, Van Oeffelen L, (2009) Iron uptake regulation in *Pseudomonas aeruginosa*. Biometals.;22(1):15-22.

Cornelissen A, Ceyssens PJ, Krylov VN, Noben J-P, Volckaert G, Lavigne R, (2012) Identification of EPS-degrading activity within the tail spikes of the novel *Pseudomonas putida* phage AF. Virology 434:251–256.

Cornelissen A, Ceyssens PJ, T'Syen J, Van Praet H, Noben JP, Shaburova OV, Krylov VN, Volckaert G, Lavigne R, (2011) The T7-related *Pseudomonas putida* phage φ15 displays virion associated biofilm degradation properties. PLoS One 6:e18597.

Cornelissen A, Hardies SC, Shaburova OV, Krylov VN, Mattheus W, Kropinski AM, Lavigne R, (2012) Complete genome sequence of the giant virus OBP and comparative genome analysis of the diverse ΦKZ-related phages. J Virol.;86(3):1844-52.

Cornelissen C N, Harvey RA, (2012), Microbiology. Lippincott Williams and Wilkins.

Coulon C, Vinogrador E, Filloux A, Sadovskaya I, (2010) Chemical analysis of cellular and extracellular carbohydrates of a biofilm forming strain *Pseudomonas aeruginosa* PA14. J Sci PLoS One;5(12):e14220.

Courtois P, (2011) *Candida* biofilms on oral biomaterials, Biomaterials-Physics and Chemistry, InTech, pp. 475-490.

Cox C, (1993) Iron and the virulence of *Pseudomonas aeruginosa*. *Pseudomonas aeruginosa*: The Opportunist.. p. 41-45.

Cullen L, McClean S, (2015) Bacterial adaptation during chronic respiratory infections. Pathogens.;4(1):66-89.

Dacheux D, Toussaint B, Richard M, Brochier G, Croize J, Attree I, (2000) *Pseudomonas aeruginosa* cystic fibrosis isolates induce rapid, type III secretion-dependent, but ExoU-independent, oncosis of macrophages and polymorphonuclear neutrophils. Infect Immun.;68: 2916–2924.

Dagan T, (2011) Phylogenomic networks. Trends Microbiol.;19: 483–491.

Dalrymple BP, Cybinski DH, Layton I, McSweeney CS, Xue GP, Swadling YJ, Lowry JB, (1997) Three *Neocallimastix patriciarum* esterases associated with the degradation of complex polysaccharides are members of a new family of hydrolases. Microbiology 143:2605-2614.

Danis-Wlodarczyk K, Olszak T, Arabski M, Wasik S, Majkowska-Skrobek G, Augustyniak D, Gula G, Briers Y, Jang HB, Vandenheuvel D, Duda KA, Lavigne R, Drulis-Kawa Z, (2015) Characterization of the newly isolated lytic bacteriophages KTN6 and KT28 and their efficacy against *Pseudomonas aeruginosa* biofilm. PLoS One.;10(5):e0127603.

Danis-Wlodarczyk K, Vandenheuvel D, Jang HB, Briers Y, Olszak T, Arabski M, Wasik S, Drabik M, Higgins G, Tyrrell J, Harvey BJ, Noben JP, Lavigne R, Drulis-Kawa Z, (2016), A proposed integrated approach for the preclinical evaluation of phage therapy in *Pseudomonas* infections. Sci Rep.;6:28115.

Darling KE, Dewar A, Evans TJ, (2004) Role of the cystic fibrosis transmembrane conductance regulator in internalization of *Pseudomonas aeruginosa* by polarized respiratory epithelial cells. Cell Microbiol.;6(6):521-33.

Dave RN, Joshi HM, Venugopalan VP, (2011) Novel biocatalytic polymer-based antimicrobial coatings as potential ureteral biomaterial: preparation and *in vitro* performance evaluation. Antimicrob Agents Chemother.;55(2):845-53.

Davidson AR, Cardarelli L, Pell LG, Radford DR, Maxwell KL, (2012) Long noncontractile tail machines of bacteriophages. Adv Exp Med Biol.;726:115-42.

Davies G, Henrissat B, (1995) Structures and mechanisms of glycosyl hydrolases. Structure 3:853–859.

Davis KM, Weiser JN, (2011) Modifications to the peptidoglycan backbone help bacteria to establish infection. Infect Immun.;79(2):562-70.

De Beer TAP, Berka K, Thornton JM, Laskowski RA, (2014) PDBsum additions. Nucleic Acids Res., 42, D292-D296.

De Smet J, Zimmermann M, Kogadeeva M, Ceyssens PJ, Vermaelen W, Blasdel B, Jang HB, Sauer U, Lavigne R, (2016) High coverage metabolomics analysis reveals phage-specific alterations to *Pseudomonas aeruginosa* physiology during infection. ISME J.;10(8):1823-35.

de Soyza A, Hall AJ, Mahenthiralingam E, Drevinek P, Kaca W, Drulis-Kawa Z, Stoitsova SR, Toth V, Coenye T, Zlosnik JE, Burns JL, Sá-Correia I, De Vos D, Pirnay JP, Kidd TJ, Reid D, Manos J, Klockgether J, Wiehlmann L, Tümmler B, McClean S, Winstanley C, (2013) Developing an international *Pseudomonas aeruginosa* reference panel. Microbiologyopen.; 2: 1010–1023.

De Vries J, Harms K, Broer I, Kriete G, Mahn A, Düring K, Wackernagel W, (1999) The bacteriolytic activity in transgenic potatoes expressing a chimeric T4 lysozyme gene and the effect of T4 lysozyme on soil- and phytopathogenic bacteria. Syst. Appl. Microbiol., 22, 280–286.

Deng Q, Barbieri JT, (2008) Molecular mechanisms of the cytotoxicity of ADP-ribosylating toxins. Annu. Rev. Microbiol. 62:271–288.

Deresinski S, (2009) Bacteriophage therapy: exploiting smaller fleas. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, *48*, 1096–1101.

Deretic V, Gill JF, Chakrabarty AM, (1987) Gene algD coding for GDP mannose dehydrogenase is transcriptionally activated in mucoid *Pseudomonas aeruginosa*. J. Bacteriol. 169:351–358.

Dietrich LEP, Price-Whelan A, Petersen A, Whiteley M, Newman DK, (2006) The phenazine pyocyanin is a terminal signalling factor in the quorum sensing network of *Pseudomonas aeruginosa*. Mol Microbiol.;61: 1308–1321.

Dong Y, Zhang LH, (2005) Quorum sensing and quorum-quenching enzymes. J. Microbiol., 43, 101-109.

Donlan RM, (2009). Preventing biofilms of clinically relevant organisms using bacteriophage. Trends Microbiol., 17(2), pp. 66–72.

Döring G, Parameswaran IG, Murphy TF, (2011) Differential adaptation of microbial pathogens to airways of patients with cystic fibrosis and chronic obstructive pulmonary disease. FEMS Microbiol. Rev. 35:124 – 146.

Dorscht J, Klumpp J, Bielmann R, Schmelcher M, Born Y, Zimmer M, Calendar R, Loessner MJ, (2009) Comparative genome analysis of *Listeria bacteriophages* reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. J Bacteriol.;191: 7206–7215.

Driscoll JA, Brody SL, Kollef MH, (2007) The epidemiology, pathogenesis and treatment of *Pseudomonas aeruginosa* infections. Drugs. 2007;67(3):351-68.

Drulis-Kawa Z, Majkowska-Skrobek G, Boczkowska B, Delattre AS, Lavigne R, (2012) Learning from bacteriophages – Advantages and limitations of phage and phage-encoded protein applications. Curr Protein Pept Sci.;13(8):699-722.

Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B, (2015) Bacteriophages and phage-derived proteins – application approaches. Current Medicinal Chemistry, *22*, 1757-1773.

Drulis-Kawa Z, Olszak T, Danis K, Majkowska-Skrobek G, Ackermann HW, (2014) A giant *Pseudomonas* phage from Poland. Arch Virol.;159(3):567-72.

Dubos R, Avery OT, (1931) Decomposition of the capsular polysaccharide of pneumococcus type III by a bacterial enzyme. J Exp Med; 54:51-71; PMID:19869902.

Dumont D, Noben JP, Raus J, Stinissen P, Robben J, (2004) Proteomic analysis of cerebrospinal fluid from multiple sclerosis patients. Proteomics 4: 2117–24.

Dzierżanowska D, (2008) Zakażenia szpitalne. Bielsko-Biała: Alfa Medica Press.

Engel J, Balachandran P, (2009) Role of *Pseudomonas aeruginosa* type III effectors in disease. Curr. Opin. Microbiol.; 12:61–66.

Engel J, Eran Y, (2012) Subversion of mucosal barrier polarity by *Pseudomonas aeruginosa*. Front. Microbiol. 2:114.

Epstein TM, Samanta U, Kirby SD, Cerasoli DM, Bahnson BJ, (2009) Crystal structures of brain group-VIII phospholipase A2 in nonaged complexes with the organophosphorusnerve agents soman and sarin. Biochemistry.;48(15):3425-35.

Erb ML, Kraemer JA, Coker JK, Chaikeeratisak V, Nonejuie P, Agard DA, Pogliano J, (2014) A bacteriophage tubulin harnesses dynamic instability to center DNA in infected cells. Elife.;3.

Erb ML, Pogliano J, (2013) Cytoskeletal proteins participate in conserved viral strategies across kingdoms of life. Curr Opin Microbiol.;16(6):786-9.

Escolar L, Pe´rez-Martı´n J, de Lorenzo V, (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. J Bacteriol 181:6223–6229.

Essoh C, Blouin Y, Loukou G, Cablanmian A, Lathro S, Kutter E, Thien HV, Vergnaud G, Pourcel C, (2013) The susceptibility of *Pseudomonas aeruginosa* strains from cystic fibrosis patients to bacteriophages. PLoS One.; (4):e60575.

Evans DJ, Frank DW, Finck-Barbancon V, Wu C, Fleiszig SM, (1998) *Pseudomonas aeruginosa* invasion and cytotoxicity are independent events, both of which involve protein tyrosine kinase activity. Infect. Immun. 66: 1453–1459.

Fauvart M, De Groote VN, Michiels J, (2011), Role of persister cells in chronic infections: clinical relevance and perspectives on anti-persister therapies. J Med Microbiol.;60(Pt 6):699-709.

Fenton M, R. Keary, O. McAuliffe, R. P. Ross, J. O'Mahony, and A. Coffey, (2013) Bacteriophage- derived peptidase CHAPk eliminates and prevents Staphylococcal biofilms. Int J Microbiol.;2013:625341.

Fenton M, Ross P, McAuliffe O, O'Mahony J, Coffey A, (2010) Recombinant bacteriophage lysins as antibacterials. Bioeng. Bugs, 1(1):9-16.

Fernebro J, (2011) Fighting bacterial infections - Future treatment options. Drug Resist. Updat., 14(2), pp. 125–139.

Filatova LY, Becker SC, Donovan DM, Gladilin AK, Klyachko NL, (2010) LysK, the enzyme lysing *Staphylococcus aureus* cells: Specific kinetic features and approaches towards stabilization. Biochimie, 92(5):507-13..

Finn RD, Clements J, Eddy SR, (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res.39, W29–W37.

Finn RD, Clements J, Eddy SR, (2011) HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Research*, 39:W29-37.

Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer EL, Bateman A, (2006) Pfam: clans, web tools and services. Nucleic Acids Res. 34, D247–D251.

Firoved AM, Deretic V, (2003) Microarray analysis of global gene expression in mucoid *Pseudomonas aeruginosa*. J. Bacteriol. 185:1071–1081.

Fischetti VA, (2010) Bacteriophage endolysins: A novel anti-infective to control Gram-positive pathogens. Int J Med Microbiol.;300(6):357-62.

Fleiszig SM, Zaidi TS, Pier GB, (1995) *Pseudomonas aeruginosa* invasion of and multiplication within corneal epithelial cells *in vitro*. Infect Immun 63, 4072–4077.

Fleiszig SM, Zaidi TS, Preston MJ, Grout M, Evans DJ, Pier GB, (1996) Relationship between cytotoxicity and corneal epithelial cell invasion by clinical isolates of *Pseudomonas aeruginosa*. Infect Immun 64: 2288– 2294.

Flemming HC, Ridgway H, (2009) Biofilm control: conventional and alternative approaches. Marine and Industrial Biofoulingm 4(1), 103-117.

Flemming HC, Wingender J, (2010) The biofilm matrix. Nat. Rev. Microbiol., 8(9):623-33.

Fokine A, Battisti AJ, Bowman VD, Efimov AV, Kurochkina LP, Chipman PR, Mesyanzhinov VV, Rossmann MG, (2007) Cryo-EM study of the *Pseudomonas* bacteriophage KZ. Structure 15:1099 –1104.

Folkesson A, Jelsbak L, Yang L, Johansen HK, Ciofu O, Hoiby N, Molin S, (2012) Adaptation of *Pseudomonas aeruginosa* to the cystic fibrosis airway: an evolutionary perspective. Nat. Rev. Microbiol.; 10:841–851.

Freiberg A, Morona R, Van Den Bosch L, Jung C, Behlke J, Carlin N, Seckler R, Baxa U, (2003) The tail spike protein of *Shigella* phage Sf6. J. Biol. Chem. *278*, 1542-1548.

Friedman L, Kolter R, (2004) Genes involved in matrix formation in *Pseudomonas aeruginosa* PA14 biofilm. J Sci Mol Microbiol;51(3):675e90.

Frirdich E, Whitfield C, (2005) Characterization of Gla(KP), a UDPgalacturonic acid C4-epimerase from *Klebsiella pneumoniae* with extended substrate specificity. J Bacteriol 187:4104–4115.

Fujitani S, Sun HY, Yu VL, Weingarten JA, (2011) Pneumonia due to *Pseudomonas aeruginosa*. Part I. Epidemiology, clinical diagnosis, and source. Chest 139:909–919. 10.1378/chest.10-0166.

Galal YS, Youssef MR, Ibrahiem SK, (2016) Ventilator-associated pneumonia: incidence, risk factors and outcome in paediatric intensive care units at Cairo University Hospital. J Clin Diagn Res. ;10(6):SC06-11.

Garbe J, Wesche A, Bunk B, Kazmierczak M, Selezska K, Rohde C, Jahn D, Schobert M, (2010) Characterization of JG024, a *Pseudomonas aeruginosa* PB1-like broad host range phage under simulated infection conditions. BMC Microbiol. BioMed Central Ltd; 10: 301.

Garron ML, Cygler M, (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. Glycobiology 20: 1547–1573.

Gasson MJ, (1991) Viral products. US patent 6,083,684A.

Gawande PV, Leung KP, Madhyastha S, (2014) Antibiofilm and antimicrobial efficacy of DispersinB®-KSL-W peptide-based wound gel against chronic wound infection associated bacteria. Curr. Microbiol., 68(5):635-41.

Gellatly SL, Hancock RE, (2013) *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. Pathog Dis.;67(3):159-73.

Ghafoor A, Hay ID, Rehm BH, (2011) Role of exopolysaccharides in *Pseudomonas aeruginosa* biofilm formation and architecture, Appl. Environ. Microbiol., 77(15):5238-46.

Glonti T, Chanishvili N, Taylor PW, (2010) Bacteriophage-derived enzyme that depolymerizes the alginic acid capsule associated with cystic fibrosis isolates of *Pseudomonas aeruginosa*. Journal of Applied Microbiology, 108(2):695-702.

Golshahi L, Lynch KH, Dennis JJ, Finlay WH, (2011) *In vitro* lung delivery of bacteriophages KS4-M and φKZ using dry powder inhalers for treatment of *Burkholderia cepacia complex* and *Pseudomonas aeruginosa* infections in cystic fibrosis. Journal of Applied Microbiology.; 110:106–117.

Górska S, Grycko P, Rybka J, Gamian A, (2007) Exopolysaccharides of lactic acid bacteria: structure and biosynthesis. Postepy Hig Med Dosw (Online);61:805-18.

Gould I, (2006) Costs of hospital-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) and its control. International J Antimicrob Ag.;28(5):379-84.

Govan JR, Deretic V, (1996) Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Microbiol.Rev.60,539–574.

Gründling A, Missiakas DM, Schneewind O, (2006) *Staphylococcus aureus* mutants with increased lysostaphin resistance. J Bacteriol.;188(17):6286-97.

Gupta R, Gupta N, Rathi P, (2004) Bacterial lipases: an overview of production, purification and biochemical properties. Appl Microbiol Biotechnol 64:763–781.

Gutiérrez D, Briers Y, Rodríguez-Rubio L, Martínez B, Rodríguez A, Lavigne R, García P, (2015) Role of the Pre-neck Appendage Protein (Dpo7) from Phage vB_SepiS-phiIPLA7 as an Anti-biofilm Agent in Staphylococcal Species.CFront Microbiol.;6:1315.

Gutierrez D, Ruas-Madiedo P, Martinez B, Rodriguez A, Garcia P, (2014) Effective removal of staphylococcal biofilms by the endolysin LysH5, PLoS One;9(9):e107307.

Haecker A, (2009) MRSA im Krankenhaus, Konsequenzen und Lösungen für das Klinikmanagement. KU Gesundheitsmanagement.;8:61-2.

Hajjar AM, Ernst RK, Tsai JH, Wilson CB, Miller SI, (2002) Human Toll-like receptor 4 recognizes host-specific LPS modifications. Nat. Immunol. 3, 354–359.

Hancock RE, Mutharia LM, Chan L, Darveau RP, Speert DP, Pier GB, (1983) *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis: a class of serum-sensitive, nontypable strains deficient in lipopolysaccharide O side chains. Infect.Immun.42,170–177.

Hanlon GW, (2007) Bacteriophages: an appraisal of their role in the treatment of bacterial infections, Int. J. Antimicrob. Agents, 30(2):118-28.

Hanlon GW, Denyer SP, Olliff CJ, Ibrahim LJ, (2001). Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through *Pseudomonas aeruginosa* biofilms. Appl. Environ. Microbiol., 67(6):2746-53.

Hargreaves KR, Kropinski AM, Clokie MR, (2014) Bacteriophage behavioral ecology: How phages alter their bacterial host's habits. Bacteriophage.;4:e29866.

Harmsen M, Yang L, Pamp SJ, Tolker-Nielsen T, (2010) An update on *Pseudomonas aeruginosa* biofilm formation, tolerance, and dispersal, FEMS Immunol. Med. Microbiol., 59(3):253-68.

Harper D, Parracho H, Walker J, Sharp R, Hughes G, Werthén M, Lehman S, Morales S, (2014) Bacteriophages and biofilms. Antibiotics 3:270–284.

Harvey J, Keenan KP, Gilmour A, (2007), Assessing biofilm formation by *Listeria monocytogenes* strains. Food Microbiol.;24(4):380-92.

Hassett DJ, Charniga L, Bean K, Ohman DE, Cohen MS, (1992) Response of *Pseudomonas aeruginosa* to pyocyanin: mechanisms of resistance, antioxidant defenses, and demonstration of a manganese-cofactored superoxide dismutase. Infect Immun.;60: 328–336.

Hassett DJ, Korfhagen TR, Irvin RT, Schurr MJ, Sauer K, Lau GW, Sutton MD, Yu H, Hoiby N, (2010) *Pseudomonas aeruginosa* biofilm infections in cystic fibrosis: Insights into pathogenic processes and treatment strategies. Exp. Opin. Ther. Targets, *14*, 117–130.

Hay ID, Gatland K, Campisano A, Jordens JZ, Rehm BH, (2009) Impact of alginate overproduction on attachment and biofilm architecture of a supermucoid *Pseudomonas aeruginosa* strain. Appl Environ Microbiol.;75(18):6022–5.

Hendrix RW, (1999) Evolution: the long evolutionary reach of viruses. Curr Biol. 16-30;9(24):R914-7.

Hendrix RW, (2009) Jumbo bacteriophages. Curr Top Microbiol Immunol.;328:229-40.

Hertveldt K, Lavigne R, Pleteneva E, Sernova N, Kurochkina L, Korchevskii R, Robben J, Mesyanxhinov V, Krylov VN, Volckaert G, (2005) Genome comparison of *Pseudomonas aeruginosa* large phages. J. Mol. Biol. 354, 536–545.

Heselpoth RD, Nelson DC, (2012) A new screening method for the directed evolution of thermostable bacteriolytic enzymes. J Vis Exp.;(69). pii: 4216.

Hogardt M, Heesemann J, (2010) Adaptation of *Pseudomonas aeruginosa* during persistence in the cystic fibrosis lung. Int J Med Microbiol.;300(8):557-62.

Hogardt, M, Hoboth C, Schmoldt S, Henke C, Bader L, Heesemann J, (2007) Stage- specific adaptation of hypermutable *Pseudomonas aeruginosa* isolates during chronic pulmonary infection in patients with cystic fibrosis. J.Infect.Dis.195, 70–80.

Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O, (2010) Antibiotic resistance of bacterial biofilms. Int J Antimicrob Agents.;35(4):322-32.

Høiby N, Ciofu O, Johansen HK , Song ZJ, Moser C, Jensen PO, Molin S, Givskov M, Tolker-Nielsen T, Bjarnsholt T, (2011) The clinical impact of bacterial biofilms, Int J Oral Sci 3: 55-65.

Holt JG, Krieg NR, Sneath PHA, Staley JT, Williams ST, (1994) Bergey's manual of determinative bacteriology (9th ed.). Lippincott Williams & Wilkins. p. 11.

Hu S, Kong J, Kong W, Guo T, Ji M, (2010) Characterization of a novel LysM domain from *Lactobacillus fermentam* bacteriophage endolysin and its use as an anchor to display heterologous proteins on the surfaces of lactic acid bacteria, Appl Environ Microbiol. 2010 Apr;76(8):2410-8.

Hübner C, Hübner NO, Kramer A, Fleßa S, (2012) Cost-analysis of PCR guided pre-emptive antibiotic treatment of *Staphylococcus aureus* infections: an analytic decision model. Eur JClin Microbiol Infect Dis.;31(11):3065-72.

Hughes KA, Sutherland IW, Clark J, Jones MV, (1998a) Bacteriophage and associated polysaccharide depolymerases-novel tools for study of bacterial biofilms. J Appl Microbiol 85:583–590.

Hughes KA, Sutherland IW, Jones MV, (1998) Biofilm susceptibility to bacteriophage attack: the role of phageborne polysaccharide depolymerase. Microbiology; 144 (Pt 11):3039-47; PMID:9846739.

Hyman P, (2012). Bacteriophages in Health and Disease. *CABI*, Wallingford.

Hyman P, Abedon ST, (2010) Bacteriophage host range and bacterial resistance. Adv Appl Microbiol; 70:217-48; PMID:20359459

Hynes WL, Walton SL, (2000) Hyaluronidases of gram-positive bacteria. FEMS Microbiol Lett 183:201–207.

Jaeger KE, Ransac S, Dijkstra BW, Colson C, van Heuvel M, Misset O, (1994) Bacterial lipases. FEMS Microbiol Rev 15:29–63.

Jaffar-Bandjee MC, Lazdunski A, Bally M, Carrere J, Chazalette JP, Galabert C, (1995) Production of elastase, exotoxin A, and alkaline protease in sputa during pulmonary exacerbation of cystic fibrosis in patients chronically infected by *Pseudomonas aeruginosa*. J. Clin. Microbiol. 33: 924 –929.

Jang HB, Fagutao FF, Nho SW, Park SB, Cha IS, Yu JE, Lee JS, Im SP, Aoki T, Jung TS, (2013) Phylogenomic network and comparative genomics reveal a diverged member of the ϕKZ-related group, marine Vibrio phage ϕJM-2012. J. Virol. 87, 12866–12878.

Jarrell K, Kropinski AM, (1977) Identification of the cell wall receptor for bacteriophage E79 in *Pseudomonas aeruginosa* strain PAO. J Virol.; 23: 461–466.

Jayaseelan S, Ramaswamy D, Dharmaraj S, (2014) Pyocyanin: production, applications, challenges and new insights. World J Microbiol Biotechnol 30 1159–1168.

Jenkins J, Pickersgill R, (2001) The architecture of parallel beta-helices and related folds. Prog. Biophys. Mol. Biol. *77*,111-115.

Jennings LK, Storek KM, Ledvina HE, Coulon C, Marmont LS, Sadovskaya I, Secor PR, Tseng BS, Scian M, Filloux A, Wozniak DJ, Howell PL, Parsek MR, (2015) Pel is a cationic exopolysaccharide that cross-links extracellular DNA in the *Pseudomonas aeruginosa* biofilm matrix. Proc Natl Acad Sci U S A.;112(36):11353-8.

Jiang Y, Chen R, Dong J, Xu Z, Gao X, (2012) Analysis of GDSL lipase (GLIP) family genes in rice (Oryza sativa), POJ 5(4):351-358.

Jiménez ER, (2009) Dextranase in sugar industry: a review. Sugar Tech 11:124–134.

Jimenez PN, Koch G, Thompson JA, Xavier KB, Cool RH, Quax WJ, (2012) The multiple signaling systems regulating virulence in P*seudomonas aeruginosa*. Microbiol. Mol. Biol. Rev. 76, 46–65.

Jokilammi A, Ollikka P, Korja M, Jakobsson E, Loimaranta V, Haataja S, Hirvonen H, Finne J, (2004) Construction of antibody mimics from a noncatalytic enzyme-detection of polysialic acid. J Immunol Methods.;295(1):149–60.

Jolivet-Gougeon A, Kovacs B, le Gall-David S, le Bars H, Bousarghin L, Bonnaure-Mallet M, Lobel B, Guille F, Soussy CJ, Tenke P, (2011) Bacterial hypermutation: Clinical implications. J. Med. Microbiol., 60, 563–573.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S, (2014) InterProScan 5: genome-scale protein function classification. Bioinformatics.;30(9):1236-40.

Juturu V, Wu JC, (2012) Microbial xylanases: engineering, production and industrial applications. Biotechnol Adv 30:1219–1227.

Kanamaru S, Ishiwata Y, Suzuki T, Rossmann MG, Arisaka F, (2005) Control of bacteriophage T4 tail lysozyme activity during the infection process. *J.Mol.Biol.* 346 p.1013.

Kang CI, Kim SH, Park WB, Lee KD, Kim HB, Kim EC, Oh MD, Choe KW, (2005) Bloodstream infections caused by antibiotic resistant gram-negative bacilli: risk factors for mortality and impact of inappropriate initial antimicrobial therapy on outcome. Antimicrob. Agents Chemother. 49:760 –766.

Kaplan JB, LoVetri K, Cardona ST, Madhyastha S, Sadovskaya I, Jabbouri S, Izano EA, (2012) Recombinant human DNase I decreases biofilm and increases antimicrobial susceptibility in staphylococci. J Antibiot, 65, 73-77.

Kegan S, (1999) Normative ethics. Wesview Press, Colorado, USA.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE, (2015) The Phyre2 web portal for protein modeling, prediction and analysis.Nat. Protoc.10, 845–858.

Ketelboeter LM, Potharla VY, Bardy SL, (2014) NTBC treatment of the pyomelanogenic *Pseudomonas aeruginosa* clinical isolatePA1111 inhibits pigment production and increases sensitivity to oxidative stress. Curr Microbiol.;69(3):343-8.

Khalifa ABH, Moissenet D, Thien DV, Khedher M, (2011) Virulence factors in *Pseudomonas aeruginosa*: mechanisms and modes of regulation. Ann Biol Clin (Paris).;69(4):393-403.

Kida Y, Higashimoto Y, Inoue H, Shimizu T, Kuwano K, (2008), A novel secreted protease from *Pseudomonas aeruginosa* activates NF-kappaB through protease-activated receptors. Cell. Microbiol. 10:1491–1504.

Kiljunen S, Hakala K, Pinta E, Huttunen S, Pluta P, Gador A, Lonnberg H, Skurnik M, (2005) Yersinophage ϕR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. Microbiology 151, 4093–4102.

Kim HS, Lee C-G, Lee EY, (2011) Alginate lyase: structure, property, and application. Biotechnol Bioprocess Eng 16:843–851.

Kim MS, Hong SS, Park K, Myung H, (2013) Genomic analysis of bacteriophage PBECO4 infecting *Escherichia coli* O157:H7. Arch Virol.;158(11):2399-403.

Kim S, Kim M, Ryu S, (2014) Development of an engineered bioluminescent reporter phage for the sensitive detection of viable *Salmonella* Typhimurium. Anal. Chem. 86, 5858–5864.

Kim S, Oh DB, Kang HA, Kwon O, (2011) Features and applications of bacterial sialidases. Appl Microbiol Biotechnol 91:1–15.

Kim SJ, Park RY, Kang SM, Choi MH, Kim CM, Shin SH, (2006) *Pseudomonas aeruginosa* alkaline protease can facilitate siderophoremediated iron-uptake via the proteolytic cleavage of transferrins. Biol. Pharm. Bull. 29:2295–2300.

Kim WS, Geider K, (2000) Characterization of a viral EPS-depolymerase, a potential tool for control of fire blight. Phytopathology 90, 1263–1268.

Kimura K, Itoh Y, (2003) Characterization of poly-gamma-glutamate hydrolase encoded by a bacteriophage genome: possible role in phage infection of *Bacillus subtilis* encapsulated with poly-gamma-glutamate. Appl Environ Microbiol 69:2491–2497.

Kirisits MJ, Parsek MR, (2006) Does *Pseudomonas aeruginosa* use intercellular signalling to build biofilm communities? Cell Microbiol.;8(12):1841-9.

Kishioka C, Okamoto K, Hassett DJ, de Mello D, Rubin BK, (1999) *Pseudomonas aeruginosa* alginate is a potent secretagogue in the isolated ferret trachea. Pediatr. Pulmonol. 27:174 –179.

Klausen M, Heydorn A, Ragas P, Lambertsen L, Aaes-Jørgensen A, Molin S, Tolker-Nielsen T, (2003) Biofilm formation by *Pseudomonas aeruginosa* wild type, flagella and type IV pili mutants. Mol. Microbiol., 48(6):1511-24.

Kobayashi H, Kobayashi O, Kawai S, (2009) Pathogenesis and clinical manifestations of chronic colonization by *Pseudomonas aeruginosa* and its biofilms in the airway tract. J Infect Chemother.;15: 125–142.

Kohanski MA, Dwyer DJ, Collins JJ, (2010) How antibiotics kill bacteria: from targets to networks. Nat Rev Microbiol.; 8(6): 423–435.

Kollef MH, Chastre J, Fagon JY, François B, Niederman MS, Rello J, Torres A, Vincent JL, Wunderink RG, Go KW, Rehm C, (2014) Global prospective epidemiologic and surveillance study of ventilator-associated pneumonia due to *Pseudomonas aeruginosa*. Crit Care Med 42:2178–2187.

Koonin EV, Dolja VV, Krupovic M, (2015) Origins and evolution of viruses of eukaryotes: The ultimate modularity. Virology.;479-480:2-25.

Kovalyova IV, Kropinski AM, (2003) The complete genomic sequence of lytic bacteriophage gh-1 infecting *Pseudomonas putida* -evidence for close relationship to the T7 group. Virology.;311(2):305-15.

Kraemer JA, Erb ML, Waddling CA, Montabana EA, Zehr EA, Wang H, Nguyen K, Pham DS, Agard DA, Pogliano J, (2012) A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. Cell.;149(7):1488-99.

Kretzer JW, Lehmann R, Schmelcher M, Banz M, Kim KP, Korn C, Loessner MJ, (2007) Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. Applied and Environmental Microbiology, 73(6):1992-2000.

Kropinski AM, (2000) Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. J. Bacteriol. *182*, 6066-6074.

Kropinski AM, (2009) Measurement of the bacteriophage inactivation kinetics with purified receptors. Methods Mol Biol.; 501: 157–160.

Krylov V, Pleteneva E, Bourkaltseva M, Shaburova O, Volckaert G, Sykilinda N, Kurochkina L, Mesyanzhinov V, (2003) *Myoviridae* bacteriophages of *Pseudomonas aeruginosa*: a long and complex evolutionary pathway. Res Microbiol.;154(4):269-75.

Krylov VN, (2014) Bacteriophages of *Pseudomonas aeruginosa*: long-term prospects for use in phage therapy. Adv Virus Res.;88:227-78.

Krylov VN, Dela Cruz DM, Hertveldt K, Ackermann HW, (2007) ‘‘φKZ-like viruses’’, a proposed new genus of myovirus bacteriophages. Arch. Virol. 152, 1955–1959.

Krylov VN, Smirnova TA, Minenkova IB, Plotnikova TG, Zhazikov IZ, Khrenova EA (1984) *Pseudomonas* bacteriophage φKZ contains an inner body in its capsid. Can. J. Microbiol. 30:758 –762.

Kumar A, Vemula PK, Ajayan PM, John G, (2008) Silver-nanoparticle-embedded antimicrobial paints based on vegetable oil. Nature Mat., 7, 236-241.

Kumar AS, Mody K, Jha B, (2007) Bacterial exopolysaccharides–a perception. J Basic Microbiol; 47:103-17; PMID:17440912.

Kumar J, (2008) Lysostaphin: an antistaphylococcal agent. Appl. Microbiol. Biotechnol. 2008, 80, 555-561.

Kurtböke I, (2012) Bacteriophages, InTech, 3-31.

Kutter E, (2005) Phage therapy: bacteriophages as natural, self-limiting antibiotics. Textbook of Natural Medicine, t.1, Philadelphia; 1147–1161.

Kutter E, (2009) Phage host range and efficiency of plating. Methods Mol Biol 501:141–149.

Laarman AJ, Bardoel BW, Ruyken M, Fernie J, Milder FJ, van Strijp JA Rooijakkers SH, (2012) *Pseudomonas aeruginosa* alkaline protease blocks complement activation via the classical and lectin pathways. J. Immunol. 188:386 –393.

Labrie SJ, Samson JE, Moineau S, (2010) Bacteriophage resistance mechanisms. Nat. Rev. Microbiol. *8*, 317-327.

Lamppa JW, Ackerman ME, Lai JI, Scanlon TC, Griswold KE, (2011) Genetically engineered alginate lyase-PEG conjugates exhibit enhanced catalytic function and reduced immunoreactivity. PLoS One, 6, e17042.

Langton Hewer SC, Smyth AR, (2014) Antibiotic strategies for eradicating *Pseudomonas aeruginosa* in people with cystic fibrosis. Cochrane Database Syst. Rev. (11): CD004197.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, (2007) ClustalW and ClustalX version 2. Bioinformatics 23(21): 2947-2948.

Lau GW, Hassett DJ, Britigan BE, (2005) Modulation of lung epithelial functions by *Pseudomonas aeruginosa*. Trends Microbiol. 13:389 –397.

Lavigne R, Briers Y, Hertveldt K, Robben J, Volckaert G, (2004) Identification and characterization of a highly thermostable bacteriophage lysozyme. Cell Mol Life Sci.;61(21):2753-9.

Lavigne R, Burkal'tseva MV, Robben J, Sykilinda NN, Kurochkina LP, Grymonprez B, Jonckx B, Krylov VN, Mesyanzhinov VV, Volckaert G, (2003) The genome of bacteriophage phiKMV, a T7-like virus infecting *Pseudomonas aeruginosa.* Virology.;312(1):49-59.

Lavigne R, Darius P, Summer EJ, Seto D, Mahadevan P, Nilsson AS, Ackermann HW, Kropinski AM, (2009) Classification of *Myoviridae* bacteriophages using protein sequence similarity. BMC Microbiol.; 9: 224.

Lawrence JG, Hatfull GF, Hendrix RW, (2002) Imbroglios of viral taxonomy: Genetic exchange and failings of phenetic approaches. J Bacteriol.;184: 4891–4905.

Lazaroaie MM, (2010) Multiple Responses of Gram-positive and Gram-negative bacteria to mixture of hydrocarbons. Braz J Microbiol.;41(3):649-67.

Le S, Yao X, Lu S, Tan Y, Rao X, Li M, Jin X, Wang J, Zhao Y, Wu NC, Lux R, He X, Shi W, Hu F, (2014) Chromosomal DNA deletion confers phage resistance to *Pseudomonas aeruginosa*. Sci Rep.;4:4738.

Lecoutere E, Ceyssens PJ, Miroshnikov KA, Mesyanzhinov VV, Krylov VN, Noben JP, Robben J, Hertveldt K, Volckaert G, Lavigne R, (2009) Identification and comparative analysis of the structural proteomes of φKZ and EL, two giant *Pseudomonas aeruginosa* bacteriophages. Proteomics 9, 3215–3219.

Lederberg J, Alexander M, Bloom BR, Hopwood DA, Hull R, Iglewski BH, Laskin AI, Oliver SG, Schaechter M, Summers WC, (2000) *Pseudomonas*. Encyclopedia of Microbiology. Second Edition. vol 3. San Diego, 876-891.

Lee DS, Kim BK, Kwon SJ, Jin HC, Park OK, (2009) Arabidopsis GDSL lipase 2 plays a role in pathogen defense via negative regulation of auxin signaling. Biochem Biophys Res Commun 379:1038-1042.

Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ, (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc.Natl.Acad.Sci.Usa 106: 4154-4159.

Leiman PG, Battisti AJ, Bowman VD, Stummeyer K, Muhlenhoff M, Gerardy-Schahn R, Scholl D, Moilineux IJ, (2007) The structures of bacteriophage K1E and K1-5 explain processive degradation of new host specificities; Journal of Molecular biology; 371; 836- 849.

Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG, (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. Cell.;118(4):419-29.

Leiman PG, Shneider MM, (2012) Contractile tail machines of bacteriophages. Adv Exp Med Biol.;726:93-114.

Lenneman BR, Rothman-Denes LB, (2015) Structural and Biochemical Investigation of Bacteriophage N4-Encoded RNA Polymerases. Biomolecules 5: 647–667.

Leplae R, Lima-Mendez G, Toussaint A, (2010) ACLAME: A classification of mobile genetic elements, update 2010. Nucleic Acids Res. 38, D57–D61.

Leroy C, Delbarre C, Ghillebaert F, Compere C, Combes D, (2008) Effects of commercial enzymes on the adhesion of a marine biofilm-forming bacterium. Biofouling, 24, 11-22.

Lewis K, (2007) Persister cells, dormancy and infectious disease. Nat. Rev. Microbiol., vol. 5, 48–56, 2007.

Lim JA, Shin H, Kang DH, Ryu S, (2012) Characterization of endolysin from a *Salmonella* Typhimurium-infecting bacteriophage SPN1S. Res Microbiol.;163(3):233-41.

Lima-Mendez G, Van Helden J, Toussaint A, Leplae R, (2008) Reticulate representation of evolutionary and functional relationships between phage genomes. Mol. Biol. Evol. 25, 762–777.

Lindberg RB, Latta RL, (1974) Phage typing of *Pseudomonas aeruginosa*: clinical and epidemiologic considerations. J Infect Dis. 130: S33-42.

Ling H, Zhao J, Zuo K, Qiu C, Yao H, Qin J, Sun X, Tang K, (2006) Isolation and expression analysis of a GDSL-like lipase gene from Brassica napus L. J Biochem Mol Biol 39:297-303.

Liu PV, (1973) Exotoxins of *Pseudomonas aeruginosa*. I. Factors that influence the production of exotoxin. A. J. Infect. Dis. 128:506 –513.

Lodise TP, Patel N, Kwa A, Graves J, Furuno JP, Graffunder E, Lomaestro B, McGregor JC, (2007) Predictors of 30-day mortality among patients with *Pseudomonas aeruginosa* bloodstream infections: impact of delayed appropriate antibiotic selection. Antimicrob. Agents Chemother. 51:3510–3515.

Loeffler JM, Djurkovic S, Fischetti VA, (2003) Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. Infect Immun.;71(11):6199-204.

Loeffler JM, Nelson D, Fischetti VA, (2001) Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science 294: 2170 –2172.

Loessner MJ, (2005) Bacteriophage endolysins - current state of research and applications. Curr Opin Microbiol.;8(4):480-7.

Longhi C, Scoarughi GL, Poggiali F, Cellini A, Carpentieri A, Seganti L, Pucci P, Amoresano A, Cocconcelli PS, Artini M, Costerton JW, Selan L, (2008) Protease treatment affects both invasion ability and biofilm formation in *Listeria monocytogenes*. Microb. Pathog., 45, 45-52.

Looijesteijn PJ, Trapet L, de Vries E, Abee T, Hugenholtz J, (2001) Physiological function of exopolysaccharides produced by *Lactococcus lactis*. Int J Food Microbiol; 64:71-80; PMID:11252513.

Łoś M, Węgrzyn G, (2012) Pseudolysogeny. Adv Virus Res.;82:339-49.

Lowe T, Eddy SR, (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.

Lu TK, Collins JJ, (2007) Dispersing biofilms with engineered enzymatic bacteriophage. Proc Natl Acad Sci U S A 104:11197–11202.

Lu TK, Koeris MS, (2011). The next generation of bacteriophage therapy. Curr Opin Microbiol.;14(5):524-31.

Lukacik P, Barnard TJ, Keller PW, Chaturvedi KS, Seddiki N, Fairman JW, Noinaj N, Kirby TL, Henderson JP, Steven AC, Hinnebusch BJ, Buchanan SK, (2012) Structural engineering of a phage lysin that targets gram-negative pathogens. Proc Natl Acad Sci U S A.;109(25):9857-62.

Lukashin AV, Borodovsky M, (1998) GeneMark.hmm: New solutions for gene finding. Nucleic Acids Res. 26, 1107–1115.

Ma L, Conover M, Lu H, Parsek MR, Bayles K, Wozniak DJ, (2009) Assembly and development of *Pseudomonas aeruginosa* biofilm matrix. PLoS Pathog;5(3):1000354.

Madigan MT, Martinko JM, Dunlap PV, Clark DP, (2008) Block biology of microorganisms (12th Edition), Cell 2, 1168.

Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL, (2012) Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect, 18(3):268-81.

Mahajan-Miklos S, Tan MW, Rahme LG, Ausubel FM, (1999) Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa-Caenorhabditis elegans* pathogenesis model. Cell.;96: 47–56.

Malloy JL, Veldhuizen RA, Thibodeaux BA, O'Callaghan RJ, Wright JR, (2005) *Pseudomonas aeruginosa* protease IV degrades surfactant proteins and inhibits surfactant host defense and biophysical functions. Am. J. Physiol. Lung Cell. Mol. Physiol. 288:L409 –L418.

Mangili A, Bica I, Snydman DR, Hamer DH, (2005) Daptomycin-resistant, methicillin-resistant *Staphylococcus aureus* bacteremia. Clin Infect Dis.;40(7):1058-60.

Mann EE, Wozniak DJ, (2012) *Pseudomonas* biofilm matrix composition and niche biology. FEMS Microbiol Rev.;36(4):893-916.

Manzanares P, Vallés S, Ramòn D, Orejas M, (2007) α-L-rhamnosidases: old and new insights. In: Polaina J, MacCabe AP (eds) Industrial enzymes. Springer Netherlands, Dordrecht, pp 117–140.

Mao J, Schmelcher M, Harty WJ, Foster-Frey J, David M, (2013) Chimeric Ply187 endolysin kills *Staphylococcus aureus* more effectively than the parental enzyme. FEMS Microbiol Lett.;342(1):30-6.

Mao YJ, Xie J, (2014) Bacteriophage polysaccharide depolymerases and biomedical applications. BioDrugs, 28(3):265-74.

Matinkhoo S, Lynch KH, Dennis JJ, Finlay WH, Vehring R, (2011) Spray-dried respirable powders containing bacteriophages for the treatment of pulmonary infections. Journal of Pharmaceutical Sciences.; 100:5197–5205.

Matsui H, Grubb BR, Tarran R, Randell SH, Gatzy JT, Davis CW, Boucher RC, (1998) Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cysticfibrosis airways disease. Cell.;95(7):1005-15.

Matsuzaki S, Inoue T, Kuroda M, Kimura S, Tanaka S, (1998) Cloning and sequencing of major capsid protein (mcp) gene of a vibriophage, KVP20, possibly related to T-even coliphages. Gene (Amst);222:25–30.

McIsaac SM, Stadnyk AW, Lin TJ, (2012) Toll-like receptors in the host defense against *Pseudomonas aeruginosa* respiratory infection and cystic fibrosis. J. Leukoc. Biol. 92:977–985.

Mena A, Smith EE, Burns JL, Speert DP, Moskowitz SM, Perez JL, Oliver A, (2008) Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. J. Bacteriol., 190, 7910–7917.

Meng X, Shi Y, Ji W, Meng X, Zhang J, Wang H, Lu C, Sun J, Yan Y, (2011) Application of a bacteriophage lysin to disrupt biofilms formed by the animal pathogen *Streptococcus suis*. Appl Environ Microbiol.;77(23):8272-9.

Merabishvili M, Pirnay JP, Verbeken G, Chanishvili N, Tediashvili M, Lashkhi N, Glonti T, Krylov V, Mast J, Van Parys L, Lavigne R, Volckaert G, Mattheus W, Verween G, De Corte P, Rose T, Jennes S, Zizi M, De Vos D, Vaneechoutte M, (2009) Quality controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. PLoS One.; 4: e4944.

Mesyanzhinov VV, Robben J, Grymonprez B, Kostyuchenko VA, Bourkaltseva MV, Sykilinda NN, Krylov VN, Volckaert G, (2002) The genome of bacteriophage φKZ of *Pseudomonas aeruginosa*. J. Mol. Biol. 317, 1–19.

Meyer JM, Neely A, Stintzi A, Georges C, Holder IA, (1996) Pyoverdin is essential for virulence of *Pseudomonas aeruginosa*. Infect Immun.;64: 518–523.

Miasnikov AN, (1997) Characterization of a novel endo-levanase and its gene from Bacillussp. L7. FEMS Microbiol Lett 154:23–28.

Micek ST, Lloyd AE, Ritchie DJ, Reichley RM, Fraser VJ, Kollef MH, (2005) *Pseudomonas aeruginosa* bloodstream infection: importance of appropriate initial antimicrobial treatment. Antimicrob. Agents Chemother. 49:1306 –1311.

Migula W, (1900) System der Bakterien. Handbuch der Morphologie, Entwicklungsgeschichte und Systematik der Bakterien. Gustav Fischer. 1900:884–885.

Miller ES, Heidelberg JF, Eisen JA, Nelson WC, Durkin AS, Ciecko A, Feldblyum TV, White O, Paulsen IT, Nierman WC, Lee J, Szczypinski B, Fraser CM, (2003).Complete genome sequence of the broad-host-range vibriophage KVP40: comparative genomics of a T4-related bacteriophage. J. Bacteriol. 185, 5220–5233.

Miller MB, Bassler BL, (2001) Quorum sensing in bacteria. Annu Rev Microbiol.;55:165-99.

Miller RB, (1996) Casiistry and modern ethics: a poetics of practival reasoning. University of Chicago Press, London, UK

Miller S. Schuler B, Seckler R, (1998) Phage P22 tail spike protein: removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. Protein Sci. *7*, 2223-2232.

Minandri F, Imperi F, Frangipani E, Bonchi C, Visaggio D, Facchini M, Pasquali P, Bragonzi A, Visca P, (2016) Role of iron uptake systems in *Pseudomonas aeruginosa* virulence and airway infection. Infect Immun.;84(8):2324-35.

Miroshnikov KA, Faizullina NM, Sykilinda NN, Mesyanzhinov VV, (2006) Properties of the endolytic transglycosylase encoded by gene 144 of *Pseudomonas aeruginosa* bacteriophage φKZ. Biochemistry (Moscow); 71:300–305.

Mishra M, Byrd MS, Sergeant S, Azad AK, Parsek MR, McPhail L, Schlesinger LS, Wozniak DJ, (2012) *Pseudomonas aeruginosa* Psl polysaccharide reduces neutrophil phagocytosis and the oxidative response by limiting complement-mediated opsonization. Cell Microbiol.;14(1):95-106.

Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK, Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD, (2015) The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. 43, D213–D221.

Mølgaard A, Kauppinen S, Larsen S, (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure 8:373-383.

Molineux IJ, (2006) The T7 group. In The Bacteriophages, R. Calendar, ed., (New York, NY: Oxford University Press), 277-301.

Monson R, Foulds I, Foweraker J, Welch M, Salmond GP, (2011) The *Pseudomonas aeruginosa* generalized transducing phage phiPA3 is a new member of the φKZ-like group of 'jumbo' phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. Microbiology.;157(Pt 3):859-67.

Montanari S, Oliver A, Salerno P, Mena A, Bertoni G, Tummler B, Cariani L, Conese M, Doring G, Bragonzi A, (2007) Biological cost of hypermutation in *Pseudomonas aeruginosa* strains from patients with cystic fibrosis. Microbiol. Sgm, *153*, 1445–1454.

Moxon ER, Kroll JS, (1990) The role of bacterial polysaccharide capsules as virulence factors. Curr Top Microbiol Immunol.;150:65-85.

Muhlenhoff M, Stummeyer K, Grove M, Sauerborn M, Gerardy-Schahn R, (2003) Proteolytic processing and oligomerization of bacteriophage-derived endosialidases. J. Biol. Chem. 278, 12634-12644.

Mulcahy LR, Burns JL, Lory S, Lewis K, (2010) Emergence of *Pseudomonas aeruginosa* strains producing high levels of persister cells in patients with cystic fibrosis. J Bacteriol 192, 6191–6199.

Muller JJ, Barbirz S, Heinle K, Freiberg A, Seckler R, Heinemann U, (2008) An inter-subunit active site between supercoiled paralllel beta-helices in the trimeric tailspike endorhamnosidase of *Shigalla flexneri* phage Sf6. Structure. *16*, 766-775.

Murakami H, Kuramoto T, Mizutani K, Nakano H, Kitahata S, (1992) Purification and some properties of a new levanase from *Bacillus sp.* No. 71. Biosci Biotechnol Biochem 56:608–613.

Murray TS, Egan M, Kazmierczak BI, (2007) *Pseudomonas aeruginosa* chronic colonization in cystic fibrosis patients. Curr Opin Pediatr.;19: 83–88.

Mushtaq N, Redpath MB, Luzio JP, Taylor PW, (2004) Prevention and cure of systemic *Escherichia coli* K1 infection by modification of the bacterial phenotype. Antimicrob. Agents Chemother. 48, 1503–1508.

Mushtaq N, Redpath MB, Luzio JP, Taylor PW, (2005) Treatment of experimental *Escherichia coli* infection with recombinant bacteriophage-derived capsule depolymerase. J Antimicrob Chemother; 56:160-5; PMID:15914489.

Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D, (2014) ARNold: A web tool for the prediction of Rho-independent transcription terminators. RNA Biol. 8, 11–13.

Nelson D, Loomis L, Fischetti VA, (2001) Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme. Proc Natl Acad Sci U S A;98(7):4107-12.

Nelson DC, Schmelcher M, Rodriguez-Rubio L, Klumpp J, Pritchard DG, Dong S, Donovan DM, (2012) Endolysins as antimicrobials. Adv Virus Res.;83:299-365.

Nicas TI, Hancock RE, (1983) *Pseudomonas aeruginosa* outer membrane permeability: isolation of a porin protein F-deficient mutant. J Bacteriol.;153(1):281-5.

Niemann H, Kwiatkowski B, Westphal U, Stirm S, (1977) Klebsiella serotype 25 capsular polysaccharide: primary structure and depolymerization by a bacteriophage-borne glycanase. J. Bacteriol. 130, 366–374.

Nowroozi J, Sepahi AA, Rashnonejad A, (2012) Pyocyanine biosynthetic genes in clinical and environmental isolates of *Pseudomonas aeruginosa* and detection of pyocyanine's Antimicrobial effects with or without colloidal silver nanoparticles. Cell J.; 14(1): 7–18.

Obeso JM, Martinez B, Rodriguez A, Garcia P, (2008) Lytic activity of the recombinant staphylococcal bacteriophage LysH5 endolysin active against *Staphylococcus aureus* in milk. Int J Food Microbiol.;128(2):212-8.

Ochoa CD, Alexeyev M, Pastukh V, Balczon R, Stevens T, (2012) *Pseudomonas aeruginosa* exotoxin Y is a promiscuous cyclase that increases endothelial tau phosphorylation and permeability. J. Biol. Chem. 287:25407–25418.

Ochsner UA, Snyder A, Vasil AI, Vasil ML, (2002) Effects of the twin arginine translocase on secretion of virulence factors, stress response, and pathogenesis. Proc. Natl. Acad. Sci. U. S. A. 99:8312–8317.

O'Flaherty S, Ross RP, Coffey A, (2009) Bacteriophage and their lysins for elimination of infectious bacteria. FEMS Microbiol Rev.;33(4):801-19.

Ohman DE, Chakrabarty AM, (1981) Genetic mapping of chromosomal determinants for the production of the exopolysaccharide alginate in a *Pseudomonas aeruginosa* cystic fibrosis isolate. Infect. Immun. 33:142–148.

Oliveira H, Thiagarajan V, Walmagh M, Sillankorva S, Lavigne R, Neves-Petersen MT, Kluskens LD, Azeredo J, (2014) A thermostable *Salmonella* phage endolysin, Lys68, with broad bactericidal properties against gram-negative pathogens in presence of weak acids. PLoS One.;9(10):e108376.

Oliveira H, Vilas Boas D, Mesnage S, Kluskens LD, Lavigne R, Sillankorva S, Secundo F, Azeredo J, (2016) Structural and enzymatic characterization of ABgp46, a novel phage endolysin with broad anti-Gram-negative bacterial activity. Front Microbiol.;7:208.

Oliver A, Canton R, Campo P, Baquero F, Blazquez J, (200) High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science, 288, 1251–1253.

Olszak T, Zarnowiec P, Kaca W, Danis-Wlodarczyk K, Augustyniak D, Drevinek P, de Soyza A, McClean S, Drulis-Kawa Z, (2015) *In vitro* and *in vivo* antibacterial activity of environmental bacteriophages against *Pseudomonas aeruginosa* strains from cystic fibrosis patients. Appl. Microbiol. Biotechnol. 99, 6021–6033.

Painter RG, Bonvillain RW, Valentine VG, Lombard GA, LaPlace SG, Nauseef WM, Wang G, (2008) The role of chloride anion and CFTR in killing of *Pseudomonas aeruginosa* by normal and CF neutrophils. J. Leukoc. Biol. 83:1345–1353.

Pajunen M, Kiljunen S, Skurnik M, (200) Bacteriophage phiYeO3-12, specific for *Yersinia enterocolitica* serotype O:3, is related to coliphages T3 and T7. J Bacteriol.; 182: 5114–5120.

Paradis-Bleau C, Cloutier I, Lemieux L, Sanschagrin F, Laroche J, Auger M, Garnier A, Levesque RC, (2007) Peptidoglycan lytic activity of the *Pseudomonas aeruginosa* phage φKZ gp144 lytic transglycosylase. FEMS Microbiol Lett.; 266:201–209.

Park PW, Pier GB, Hinkes MT, Bernfield M, (2001) Exploitation of syndecan-1 shedding by *Pseudomonas aeruginosa* enhances virulence. Nature 411:98 –102.

Park Y, Lim JA, Kong M, Ryu S, Rhee S, (2014) Structure of bacteriophage SPN1S endolysin reveals an unusual two-module fold for the peptidoglycan lytic and binding activity. Mol.Microbiol. 92: 316-325.

Parks QM, Young RL, Poch KR, Malcolm KC, Vasil ML, Nick JA, (2009), Neutrophil enhancement of *Pseudomonas aeruginosa* biofilm development: human F-actin and DNA as targets for therapy. J Med Microbiol.;58(Pt 4):492-502.

Parsek MR, Tolker-Nielsen T, (2008) Pattern formation in *Pseudomonas aeruginosa* biofilms. Curr. Opin. Microbiol., vol. 11, 560–566.

Pastagia M, Euler C, Chahales P, Fuentes-Duculan J, Krueger JG, Fischetti VA, (2011) A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive *Staphylococcus aureus* strains. Antimicrob Agents Chemother.;55(2):738-44.

Paul VD, Rajagopalan SS, Sundarrajan S, George SE, Asrani JY, Pillai R, Chikkamadaiah R, Durgaiah M, Sriram B, Padmanabhan S, (2011) A novel bacteriophage tail-associated muralytic enzyme (TAME) from phage K and its development into a potent antistaphylococcal protein. BMC Microbiol.;11:226.

Pelkonen S, Aalto J, Finne J, (1992) Differential activities of bacteriophage depolymerase on bacterial polysaccharide: binding is essential but degradation is inhibitory in phage infection of Kl-Defective *Escherichia coli*, Journal of Bacteriology, 174(23):7757-61.

Pietilä MK, Demina TA, Atanasova NS, Oksanen HM, Bamford DH, (2014) Archaeal viruses and bacteriophages: comparisons and contrasts. Trends Microbiol.;22(6):334-44.

Pires DP, Oliveira H, Melo LD, Sillankorva S, Azeredo J, (2016) Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl Microbiol Biotechnol.;100(5):2141-51.

Pirnay JP, De Vos D, Cochez C, Bilocq F, Pirson J, Struelens M, Duinslaeger L, Cornelis P, Zizi M, Vanderkelen A, (2003) Molecular epidemiology of *Pseudomonas aeruginosa* colonization in a burn unit: persistence of a multidrug-resistant clone and a silver sulfadiazine-resistant clone. J Clin Microbiol.;41(3):1192-202.

Pirnay JP, De Vos D, Cochez C, Bilocq F, Vanderkelen A, Zizi M, Ghysels B, Cornelis P, (2002) *Pseudomonas aeruginosa* displays an epidemic population structure. Environ Microbiol.; 4: 898–911.

Plotkowski MC, de Bentzmann S, Pereira SH, Zahm JM, Bajolet-Laudinat O, Roger P, Puchelle E, (1999) *Pseudomonas aeruginosa* internalization by human epithelial respiratory cells depends on cell differentiation, polarity, and junctional complex integrity. Am J Respir Cell Mol Biol 20: 880–890.

Poole K, (2011) *Pseudomonas aeruginosa*: resistance to the max. Frontiers in Microbiology 2: 1-13.

Proctor VA, Cunningham FE, Fung DYC, (1988) The chemistry of lysozyme and its use as a food preservative and a pharmaceutical. CRC Crit. Rev. Food Sci. Nutr., 26, 359-395.

Proetzel G, Ebersbach H, Walker J, (2012) Antibody Methods and Protocols, vol. 531, no. 1. Humana Press.

Qaisar U, Kruczek CJ, Azeem M, Javaid N, Colmer-Hamood JA, Hamood AN, (2016) The *Pseudomonas aeruginosa* extracellular secondary metabolite, Paerucumarin, chelates iron and is not localized to extracellular membrane vesicles. J Microbiol.;54(8):573-81.

Raaij MJ, (2008) Structure of the receptor-binding protein of bacteriophage Det7: a podoviral tail spike in a myovirus. J. Virol. 82, 2265-2273.

Rada B, Leto TL, (2013) Pyocyanin effects on respiratory epithelium: relevance in *Pseudomonas aeruginosa* airway infections. Trends Microbiol. 21:73–81.

Rajaure M, Berry J, Kongari R, Cahill J, Young R, (2015). Membrane fusion during phage lysis. Proc. Natl. Acad. Sci. *U. S. A.*, 112(17), 5497–5502.

Ramos JL, (2004) *Pseudomonas*, vols 1-4, New York: Kluwer Academic/ Plenum Publishers.

Rashel M, Uchiyama J, Ujihara T, Uehara Y, Kuramoto S, Sugihara S, Yagyu K, Muraoka A, Sugai M, Hiramatsu K, Honke K, Matsuzaki S, (2007) Efficient elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from bacteriophagephi MR11. J Infect Dis.;196(8):1237-47. Epub 2007 Sep 11.

Read RC, Roberts P, Munro N, Rutman A, Hastie A, Shryock T, Hall R, McDonald-Gibson W, Lund V, Taylor G, *et al.*, (1992) Effect of *Pseudomonas aeruginosa* rhamnolipids on mucociliary transport and ciliary beating. J. Appl. Physiol. 72:2271–2277.

Rehm BHA, (2009) *Pseudomonas*: Model Organism, Pathogen, Cell Factory, Wiley-Blackwell, 1-246.

Remmert M, Biegert A, Hauser A, Söding J, (2011) HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods. 9(2):173-5.

Resch G, Kulik EM, Dietrich FS, Meyer J, (2004) Complete genomic nucleotide sequence of the temperate bacteriophage Aa Phi 23 of *Actinobacillus actinomycetemcomitans*. J Bacteriol.;186: 5523–5528.

Rice LB, (2010) Progress and challenges in implementing the research on ESKAPE pathogens. Infect Control Hosp Epidemiol. 31 Suppl 1:S7-10.

Rice P, Longden I, Bleasby A, (2000) EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet 16:276–277.

Roach DR, Donovan DM, (2015) Antimicrobial bacteriophage-derived proteins and therapeutic applications. Bacteriophage.;5(3):e1062590.

Robakis NK, Palleroni NJ, Despreaux CW, Boublik M, Baker CA, Churn PJ, Claus GW, (1985) Isolation and characterization of two phages for Gluconobacter oxydans. J Gen Microbiol 131:2467–2473.

Rodriguez-Rojas A, Oliver A, Blazquez, J (2012) Intrinsic and environmental mutagenesis drive diversification and persistence of *Pseudomonas aeruginosa* in chronic lung infections. J. Infect. Dis., 205, 121–127.

Rodríguez-Rubio L, Gerstmans H, Thorpe S, Mesnage S, Lavigne R, Briers Y, (2016) DUF3380 domain from a *Salmonella* phage endolysin shows potent N -acetylmuramidase activity. Appl Environ Microbiol.;82(16):4975-81.

Rodríguez-Rubio L, Gutiérrez D, Donovan DM, Martínez B, Rodríguez A, García P, (2016) Phage lytic proteins: biotechnological applications beyond clinical antimicrobials. Crit Rev Biotechnol.;36(3):542-52.

Rodríguez-Rubio L, Martínez B, Donovan DM, García P, Rodríguez A, (2013) Potential of the virion-associated peptidoglycan hydrolase HydH5 and its derivative fusion proteins in milk biopreservation. PLoS One.;8(1):e54828.

Rodríguez-Rubio L, Martínez B, Donovan DM, Rodríguez A, García P, (2013) Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. Crit Rev Microbiol.;39(4):427-34.

Rodríguez-Rubio L, Martínez B, Rodríguez A, Donovan DM, García P, (2012) Enhanced staphylolytic activity of the *Staphylococcus aureus* bacteriophage vB_SauS-phiIPLA88 HydH5virion-associated peptidoglycan hydrolase: fusions, deletions, and synergy with LysH5. Appl Environ Microbiol.;78(7):2241-8.

Rodríguez-Rubio L, Martínez B, Rodríguez A, Donovan DM, Götz F, García P, (2013) The phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB_SauS-phiIPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance. PLoS One.;8(5):e64671.

Rodríguez-Rubio L, Martínez B, Zhou Y, Rodríguez A, Donovan DM, García P, (2011) Lytic activity of the virion-associated peptidoglycan hydrolase HydH5 of *Staphylococcus aureus* bacteriophage vB_SauS-phiIPLA88. BMC Microbiol. 2011 Jun 17;11:138.

Rohwer F, Youle M, (2011) Consider something viral in your research, Nat Rev Microbiol 9: 308-309.

Roux S, Hallam SJ, Woyke T, Sullivan MB, (2015) Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. Elife. 4.

Sabouri Ghannad M, Mohammadi A, (2012) Bacteriophage: time to re-evaluate the potential of phage therapy as a promising agent to control multidrug-resistant bacteria. Iran J Basic Med Sci.;15(2):693-701.

Sakuragi Y, Kolter R, (2007) Quorum sensing regulation of the biofilm matrix genes (pel) of *Pseudomonas aeruginosa*. J Bacteriol;189:5383e6.

Sambrook J, Russel DW, (2001) Molecular cloning: A laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press., New York, USA.

Santos SB, Carvalho CM, Sillankorva S, Nicolau A, Ferreira EC, Azeredo J, (2009) The use of antibiotics to improve phage detection and enumeration by the double-layer agar technique. BMC Microbiol 9:148.

Sanz J, Garcia J, Laynez J, Usobiaga P, Menendez M, (1993) Thermal stability and cooperative domains of CPL1 lysozyme and its NH2- and COOH-terminal modules: Dependence on choline binding. J Biol Chem.;268(9):6125-30.

Savalia D, Westblade LF, Goel M, Florens L, Kemp P, Akulenko N, Pavlova O, Padovan JC, Chait BT, Washburn MP, Ackermann HW, Mushegian A,Gabisonia T, Molineux I, Severinov K, (2008) Genomic and proteomic analysis of phiEco32, a novel *Escherichia coli* bacteriophage. J Mol Biol.;377(3):774-89.

Sayers EW, Barrett T, Benson D, Bolton E, Bryant S, Canese K, Chetvernin V, Church D, DiCuccio M, Federhen S, Feolo M, Fingerman I, Geer L, Helmberg W, Kapustin Y, Landsman D, Lipman D, Lu Z, Madden T, Madej T, Maglott D, Marchler-Bauer A, Miller V, Mizrachi I, Ostell J, Panchenko A, Phan L, Pruitt K, Schuler G, Sequeira E, Sherry S, Shumway M, Sirotkin K, Slotta D, Souvorov A, G S, Tatusova T, Wagner L, Wang Y, Wilbur W, Yaschenko E, Ye J, (2011) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 39:D38–51.

Schlegel HG, (2008) Mikrobiologia ogólna, Warszawa 2, 71-80.

Schleifer KH, Kandler O, (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev.;36(4):407-77.

Schmelcher M, Donovan DM, Loessner MJ, (2012) Bacteriophage endolysins as novel antimicrobials. Future Microbiol.;7(10):1147-71.

Schmelcher M, Loessner MJ, (2014) Application of bacteriophages for detection of foodborne pathogens. Bacteriophage 4, e28137.

Schmelcher M, Powell AM, Becker SC, Camp MJ, Donovan DM, (2012) Chimeric phage lysins act synergistically with lysostaphin to kill mastitiscausing *Staphylococcus aureus* in murine mammary glands. Appl. Environ. Microbiol. 78:2297–2305.

Schmiel DH, Miller VL, (1999) Bacterial phospholipases and pathogenesis. Microbes Infect. 1:1103–1112.

Scholl D, Rogers S, Adhya S, Merril CR, (2001) Bacteriophage K1-5 encodes two different tail fiber proteins, allowing it to infect and replicate on both K1 and K5 strains of *Escherichia coli*. J Virol 75:2509–2515.

Schroeder TH, Reiniger N, Meluleni G, Grout M, Coleman FT, Pier GB, (2001) Transgenic cystic fibrosis mice exhibit reduced early clearance of *Pseudomonas aeruginosa* from the respiratory tract. J. Immunol. 166, 7410–7418.

Schröter J (1872) Über einige durch bacterien gebildete *Pigmente.* In: Cohn, Beitrage zur Biologie der Pflanzen, 1, Heft 2, Breslau: Max Muller, p. 122-123.

Schuch R, Nelson D, Fischetti VA, (2002) A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature; 418:884-9; PMID:12192412.

Schwarzer D, Stunmeyer K, Gerardy-Schahn R, Muhlenhoff M, (2007) Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber protein. J. Biol. Chem. *282*, 2821-2831.

Scorpio A, Chabot DJ, Day WA, Hoover TA, Friedlander AM, (2010) Capsule depolymerase overexpression reduces *Bacillus anthracis* virulence. Microbiology.;156(Pt 5):1459-67.

Scorpio A, Tobery SA, Ribot WJ, Friedlander AM, (2008) Treatment of experimental anthrax with recombinant capsule depolymerase. Antimicrob. Agents Chemother. 52, 1014–1020.

Serwer P, Hayes SJ, Thomas JA, Hardies SC, (2007) Propagating the missing bacteriophages: a large bacteriophage in a new class, Virol J.;4:21.

Serwer P, Wright ET, Hakala KW, Weintraub ST, (2008) Evidence for bacteriophage T7 tail extension during DNA injection. BMC Res Notes.;1:36.

Shevchenko A, Wilm M, Vorm O, Mann M, (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68: 850–8.

Shikongo-Nambabi M, (2011) Control of bacterial contamination during marine fish processing. Journal of Biology and Life Science, 3 (1), 1-17.

Sieiro C, García-Fraga B, López-Seijas J, da Silva AF, Villa TG, (2012) Microbial pectic enzymes in the food and wine kndustry. In: Valdez B (ed) Food Industrial Processes-Methods and Equipment. InTech, 201–218.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG, (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using ClustalOmega. Mol Syst Biol.;7:539.

Silby MW, Winstanley C, Godfrey SA, Levy SB, Jackson RW, (2011) *Pseudomonas genomes*: diverse and adaptable. FEMS Microbiol. Rev. 35, 652–680.

Šimoliūnas E, Kaliniene L, Truncaite L, Klausa V, Zajančkauskaite A, Meškys R, (2012) Genome of *Klebsiella sp.*-infecting bacteriophage vB_KleM_RaK2. J Virol.;86(9):5406.

Skillman LC, Sutherland IW, Jones MV, (1999) The role of exopolysaccharides in dual species biofilm development. J.Appl. Microbiol. 85, 13-18.

Slonczewski JL, Foster JW, (2011). Microbiology, an evolving science (2nd ed.). W.W. Norton & Company Inc., New York.

Slonczewski JL, Foster JW, Microbiology, an evolving science, 2nd ed. 2011.

Söding J, Biegert A, Lupas AN, (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 33, W244-8.

Sofos JN, Geornaras I, (2010) Overview of current meat hygiene and safety risks and summary of recent studies on biofilms, and control of *Escherichia coli* O157:H7 in non-intact, and *Listeria monocytogenes* in ready-to-eat, meat products. Meat Sci.;86(1):2-14.

Sousa AM, Pereira MO, (2014) *Pseudomonas aeruginosa* diversification during infection development in cystic fibrosis lungs. Pathogens.;3(3):680-703.

Srey S, Jahid IK, Ha SD, (2013) Biofilm formation in food industries: A food safety concern, Food Control 31(2):572-585**.**

Starkey M, Hickman JH, Ma L, Zhang N, De Long S, Hinz A, Palacios S, Manoil C, Kirisits MJ, Starner TD, Wozniak DJ, Harwood CS, Parsek MR, (2009) *Pseudomonas aeruginosa* rugose small-colony variants have adaptations that likely promote persistence in the cystic fibrosis lung. J Bacteriol.;191(11):3492-503.

Stemberk V, Jones RP, Moroz O, Atkin KE, Edwards AM, Turkenburg JP, Leech AP, Massey RC, Potts JR, (2014) Evidence for steric regulation of fibrinogen binding to *Staphylococcus aureus* fibronectin-bindingprotein A (FnBPA). J Biol Chem.;289(18):12842-51.

Steven AC, Trus BL, Maizel JV, Unser M, Parry DA, Wall JS, Hainfeld JF, Studier FW, (1988) Molecular substructure of a viral receptor-recognition protein. The gp17 tail fiber of bacteriophage T7. J. Mol. Biol. 200, 351-365.

Stewart PS, Franklin MJ, (2008), Physiological heterogeneity in biofilms. Nat. Rev. Microbiol.,6, 199–210, 2008.

Stone R, (2002) Bacteriophage therapy: Stalin's forgotten cure. Science; 298:728–731.

Stothard P, Wishart DS, (2005), Circular genome visualization and exploration using CGView. Bioinformatics.;21(4):537-9.

Strateva T, Yordanov D, (2009) *Pseudomonas aeruginosa* - a phenomenon of bacterial resistance. J Med Microbiol.;58(Pt 9):1133-48.

Stummeyer K, Dickmanns A, Muhlenhoff M, Gerardy-Schahn R, Ficner R, (2005) Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. Nat. Struct. Mol. Biol. 12, 90-96.

Sullivan MB, Huang KH, Ignacio-Espinoza JC, Berlin AM, Kelly L, Weigele PR, DeFrancesco AS, Kern SE, Thompson LR, Young S, Yandava C, Fu R, Krastins B, Chase M, Sarracino D, Osburne MS, Henn MR, Chisholm SW, (2010) Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environ Microbiol.;12(11):3035-56.

Sun Q, Kuty GF, Arockiasamy A, Xu M, Young R, Sacchettini JC, (2009) Regulation of a muralytic enzyme by dynamic membrane topology. Nat Struct Mol Biol.;16(11):1192-4.

Sunna A, Antranikian G, (1997) Xylanolytic enzymes from fungi and bacteria. Crit Rev Biotechnol 17:39–67.

Sutherland IW, (1999) Polysaccharases for microbial exopolysaccharides. Carbohydrate Polymers; 38:319-28.

Sutherland IW, (2001) The biofilm matrix - An immobilized but dynamic microbial environment. Trends Microbiol., 9, (5), 222–227.

Sutherland IW, (2001). Biofilm exopolysaccharides: A strong and sticky framework. Microbiology,147, 3–9.

Sutherland IW, Hughes KA, Skillman LC, Tait K, (2004) The interaction of phage and biofilms. FEMS Microbiol. Lett., 232 (1), 1–6.

Sycheva LV, Shneider MM, Sykilinda NN, Ivanova MA, Miroshnikov KA, Leiman PG, (2012) Crystal structure and location of gp131 in the bacteriophage φKZ virion. Virology.;434(2):257-64.

Sykilinda NN, Bondar AA, Gorshkova AS, Kurochkina LP, Kulikov EE, Shneider MM, Kadykov VA, Solovjeva NV, Kabilov MR, Mesyanzhinov VV, Vlassov VV, Drukker VV, Miroshnikov KA, (2014) Complete genome sequence of the novel giant *Pseudomonas* phage PaBG. Genome Announc.;2(1). pii: e00929-13.

Szewczyk EM, (2005) Diagnostyka bakteriologiczna, PWN, 102-103.

Tait K, Skillman LC, Sutherland IW, (2002) The efficacy of bacteriophage as a method of biofilm eradication. Biofouling 18:305–311.

Taylor PK, Yeung AT, Hancock RE, (2014) Antibiotic resistance in *Pseudomonas aeruginosa* biofilms: towards the development of novel anti-biofilmtherapies. J Biotechnol.10;191:121-30.

Terada LS, Johansen KA, Nowbar S, Vasil AI, Vasil ML, (1999) *Pseudomonas aeruginosa* hemolytic phospholipase C suppresses neutrophil respiratory burst activity. Infect. Immun. 67:2371–2376.

Thallinger B, Prasetyo EN, Nyanhongo GS, Guebitz GM, (2013) Antimicrobial enzymes: an emerging strategy to fight microbes and microbial biofilms. Biotechnol J.;8(1):97-109.

Thomas JA, Rolando MR, Carroll CA, Shen PS, Belnap DM, Weintraub ST, Serwer P, Hardies SC, (2008) Characterization of *Pseudomonas chlororaphis* myovirus 2012-1 via genomic sequencing, mass spectrometry, and electron microscopy. Virology 376:330 –338.

Thunnissen AM, Rozeboom HJ, Kalk KH, Dijkstra BW, (1995) Structure of the 70-kDa soluble lytic transglycosylase complexed with bulgecin A. Implications for theenzymatic mechanism. Biochemistry.;34(39):12729-37.

Tiina M, Sandholm M, (1989) Antibacterial effect of the glucose oxidase-glucose system on food-poisoning organisms. Int. J. Food Microbiol., 8, 165-174.

Tolba M, Ahmed MU, Tlili C, Eichenseher F, Loessner MJ, Zourob M, (2012) A bacteriophage endolysin-based electrochemical impedance biosensor for the rapid detection of *Listeria* cells. Analyst, 137(24):5749-56.

Towar K, Lectures in Microbiology, University of Wisconsin-Madison [Online] http://textbookofbacteriology.net/themicrobialworld/Phage.html.

Trautmann M, Lepper PM, Haller M, (2005) Ecology of *Pseudomonas aeruginosa* in the intensive care unit and the evolving role of water outlets as a reservoir of the organism. Am J Infect Control.;33: S41–49.

Ueno H, Yonesaki T, (2004) Phage-induced change in the stability of mRNAs. Virology 329: 134–141.

UniProt Consortium (2015), UniProt: a hub for protein information. Nucleic Acids Res.;43.

Updegraff EP, Zhao F, Preuss D, (2009) The extracellular lipase EXL4 is required for efficient hydration of *Arabidopsis pollen*. Sexual plant reproduction 22:197-204.

Upton C, Buckley JT, (1995) A new family of lipolytic enzymes? Trends Biochem Sci 20:178-179.

Uzan M, (2009) RNA processing and decay in bacteriophage T4. Prog Mol Biol Transl Sci 85: 43–89.

Vaara M, (1992) Agents that increase the permeability of the outer membrane. Microbiol Rev. 56(3):395-411.

Van den Bossche A, Ceyssens PJ, De Smet J, Hendrix H, Bellon H, Leimer N, Wagemans J, Delattre AS, Cenens W, Aertsen A, Landuyt B, Minakhin L, Severinov K, Noben JP, Lavigne R, (2014) Systematic identification of hypothetical bacteriophage proteins targeting key protein complexes of *Pseudomonas aeruginosa*. J. Proteome Res.13, 4446–4456.

Varea J, Monterroso B, Saiiz J, Lopez-Zumel C, Garcia J, Laynez J, Garcia P, Mendez M, (2004) Structural and thermodynamic characterization of Pal, a phage natural chimeric lysin active against pneumococci. J Biol Chem.;279(42):43697-707.

Verma V, Harjai K, Chhibber S, (2010) Structural changes induced by a lytic bacteriophage make ciprofloxacin effective against older biofilm of *Klebsiella pneumoniae*. Biofouling 26, 729–737.

Vollmer W, (2008) Structural variation in the glycan strands of bacterial peptidoglycan. FEMS Microbiol Rev.;32(2):287-306.

Wagemans J, Blasdel BG, Van den Bossche A, Uytterhoeven B, De Smet J, Paeshuyse J, Cenens W, Aertsen A, Uetz P, Delattre AS, Ceyssens PJ, Lavigne R, (2014) Functional elucidation of antibacterial phage ORFans targeting *Pseudomonas aeruginosa*. Cell Microbiol.;16(12):1822-35.

Wagner VE, Iglewski BH, (2008) *P. aeruginosa* biofilms in CF infection. Clin Rev Allergy Immunol.;35(3):124-34.

Walmagh M, Boczkowska B, Grymonprez B, Briers Y, Drulis-Kawa Z, Lavigne R, (2013) Characterization of five novel endolysins from Gram-negative infecting bacteriophages. Appl.Microbiol. Biotechnol., 97, (10), 4369–4375.

Walmagh M, Briers Y, dos Santos SB, Azeredo J, Lavigne R, (2012) Characterization of modular bacteriophage endolysins from *Myoviridae* phages OBP, 201φ2-1 and PVP-SE1. PLoS One.;7(5):e36991.

Walter M, Fiedler C, Grassl R, Biebl M, Rachel R, Hermo-Parrado XL, Llamas-Saiz AL, Seckler R, Miller S, van Raaij MJ, (2008) Structure of the receptor-binding protein of bacteriophage Det7: a podoviral tail spike in a myovirus. J. Virol. *82*, 2265-2273.

Wang IN, (2006) Lysis timing and bacteripophage fitness. Genetics 172, 17-26.

Wang IN, Smith DL, Young R, (2000) Holins: the protein clocks of bacteriophage infections. Annu Rev Microbiol.;54:799-825.

Wang J, Pantopoulos K, (2011) Regulation of cellular iron metabolism. Biochem. J. 434, 365–381.

Waseh S, Hanifi-Moghaddam P, Coleman R, Masotti M, Ryan S, Foss M, MacKenzie R, Henry M, Szymanski CM, Tanha J, (2010) Orally administered P22 phage tail spike protein reduces salmonella colonization in chickens: prospects of a novel therapy against bacterial infections. PLoS One; 5:e13904; PMID:21124920.

Wass MN, Kelley LA, Sternberg MJ, (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res.;38.

Wei Q, Ma LZ, (2013) Biofilm matrix and its regulation in *Pseudomonas aeruginosa*. Int J Mol Sci.;14: 20983–21005.

Westbye A, Fogg P, Beatty J, (2014) Endolysin expression, purification and activity determination by zymography. [Online]. Available: http://www.bioReferences84protocol.org/e1208.

Westbye AB, Leung MM, Florizone SM, Taylor TA, Johnson JA, Fogg PC, Beatty JT, (2013) Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the *Rhodobacter capsulatus* gene transfer agent. J Bacteriol.;195(22):5025-40.

Westphal O, Jann K, (1965) Methods in Carbohydrate Chemistry. In: Whistler LR, editor. Methods in Carbohydrate Chemistry Volume V. New York: Academic Press; 83.

Williams BJ, Dehnbostel J, Blackwell TS, (2010) *Pseudomonas aeruginosa*: host defence in lung diseases. Respirology 15:1037–1056.

Winstanley C, Fothergill JL, (2009) The role of quorum sensing in chronic cystic fibrosis *Pseudomonas aeruginosa* infections. FEMS Microbiol. Lett. 290, 1–9.

Winstanley C, O'Brien S, Brockhurst MA, (2016) *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. Trends Microbiol.;24(5):327-37.

Wohlkönig A, Huet J, Looze Y, Wintjens R, (2010) Structural relationship in lysozyme superfamily: significant evidence of glycoside hydrolase signature motifs. PLoS One.;5(11):e15388.

Wolf P, Elsasser-Beile U, (2009) *Pseudomonas* exotoxin A: from virulence factor to anti-cancer agent. Int. J. Med. Microbiol. 299:161–176.

Wolska K, Kot B, Piechota M, Frankowska A, (2013) Oporność *Pseudomonas aeruginosa* na antybiotyki. Postępy Hig Med Dosw, *67*, 1300–1311.

Wong TY, Preston LA, Schiller NL, (2000) Alginate lyase: review of major sources and enzyme characteristics, structure-function analysis, biological roles, and applications. Annu Rev Microbiol 54:289– 340.

Worlitzsch D, Tarran R, Ulrich M, Schwab U, Cekici A, Meyer KC, Birrer P, Bellon G, Berger J, Weiss T, Botzenhart K, Yankaskas JR, Randell S, Boucher RC, Döring G, (2002) Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. J Clin Invest.;109(3):317-25.

Wróblewska M, (2006) Novel therapies of multidrug-resistant *Pseudomonas aeruginosa* and *Acinetobacter spp*. infections: the state of the art, Arch Immunol Ther Exp (Warsz);54(2):113-20.

Xiang Y, Leiman PG, Li L, Grimes S, Anderson DL, Rossmann MG, (2009) Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. Mol. Cell. 34, 375-386.

Yadav V, Yadav PK, Yadav S, Yadav KDS, (2010) α-l-Rhamnosidase: a review. Process Biochem 45:1226–1235.

Yakunina M, Artamonova T, Borukhov S, Makarova KS, Severinov K, Minakhin L, (2015) A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage. Nucleic Acids Res.2;43(21):10411-20.

Yamada T, Satoh S, Ishikawa H, Fujiwara A, Kawasaki T, Fujie M, Ogata H, (2010) A jumbo phage infecting the phytopathogen *Ralstonia solanacearum* defines a new lineage of the Myoviridae family. Virology.;398(1):135-47.

Yan J, Mao J, Xie J, (2014) Bacteriophage polysaccharide depolymerases and biomedical applications. *BioDrugs* 28 265–274.

Yoshimura K, Toibana A, Nakahama K, (1988) Human lysozyme: sequencing of a cDNA, and expression and secretion by *Saccharomyces cerevisiae*. Biochem Biophys Res Commun.;150(2):794-801.

Young R, (2014) Phage lysis: three steps, three choices, one outcome. J Microbiol.,52(3), 243–258.

Yu H, Head NE, (2002) Persistent infections and immunity in cystic fibrosis, Frontiers in Bioscience 7, d442-457.

Zabner J, Karp P, Seiler M, Phillips SL, Mitchell CJ, Saavedra M, Welsh M, Klingelhutz AJ, (2003) Development of cystic fibrosis and noncystic fibrosis airway cell lines. Am. J. Physiol. - Lung Cell. Mol. Physiol. 284, L844–L854.

Zegans ME, Wagner JC, Cady KC, Murphy DM, Hammond JH, O'Toole GA, (2009) Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. J Bacteriol.;191(1):210-9.

Supplementary data

Table 4.1: The lytic activity of KT28, KTN6, KTN4, PA5oct towards *P. aeruginosa* Pirnay's panel. The comparisement with host range of representatives of *Luz7virus* (LUZ7), *Lit1virus* (LIT1), *LUZ24virus* (LUZ24), *Phikmvvirus* (LUZ19, LKD16, LKA1, φKMV), *Pbunavirus* (LBL3, LMA2, LSL4, KT28, KTN6), *Phikzvirus* (KTN4, φKZ) **(**Pirnay *et al.*, 2002). Grey box –active, white box – no activity.

Strain \ Phage	φKZ	KTN4	PA5oct	KTN6	KT28	φKMV	LKD16	LUZ19	LKA1	LIT1	LUZ7	LUZ24	LSL4	LMA2	LBL3
US449															
LMG14083															
Bu007															
PAO23															
Aa249															
US448															
PAO1Krylov															
Lo050															
US450															
Li004															
Be128															
Lo053															
ATCC27853															
Br906															
PhDW6															
So099															
Aa245															
LMG5031															
C17															
C19															
Lo049															
Br642															
Li012															
Is579															
C															
Lw1047															
Br257															
Br667															

Strain \ Phage	φKZ	KTN4	PA5oct	KTN6	KT28	φKMV	LKD16	LUZ19	LKA1	LIT1	LUZ7	LUZ24	LSL4	LMA2	LBL3
C18															
LMG2107															
Bu004															
C1															
PAO29															
C13															
C2															
Aa246															
LMG14084															
Br735															
Mi162															
SG17M															
Pr335															
Is580															
Li009															
Be136															
Is573															
Br908															
TuD199															
So095															
So092															
Br229															
SG50M															
PT31M															
Mi151															
US447															
Bo548															
Br680															
PA6															
Summary	21	20	15	24	16	21	22	23	3	10	18	16	10	12	26
%	37%	35%	26%	42%	28%	37%	39%	40%	5%	17%	32%	28%	17%	21%	46%

Table S 4.2: The *Pseudomonas* phage KT28 genome annotation.

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp001	Uncharacterized protein	1	303	-	ATG	TAA	303	100	*Pseudomonas* phage NH-4	4.7e-58
gp002	Uncharacterized protein	315	461	-	ATG	TAA	147	48	*Pseudomonas* phage NH-4	8.0e-23
gp003	Uncharacterized protein	520	819	-	ATG	TGA	300	100	KT28 unique protein	
gp004	Uncharacterized protein	865	1059	-	ATG	TAA	195	64	*Pseudomonas* phage NH-4	8.1e-36
gp005	Terminase large subunit	1227	2609	+	ATG	TAA	1383	460	*Pseudomonas* phage NH-4	2.3e-314
gp006	Uncharacterized protein	2646	3029	-	ATG	TAA	384	127	*Pseudomonas* phage KPP12	5.3e-75
gp007	Uncharacterized protein	3029	3244	-	ATG	TAA	216	71	*Pseudomonas* phage SN	2.1e-41
gp008	Uncharacterized protein	3244	3594	-	ATG	TGA	351	116	*Pseudomonas* phage PB1	1.9e-69
gp009	Uncharacterized protein	3639	4073	-	ATG	TGA	435	144	*Pseudomonas* phage PB1	7.6e-78
gp010	Uncharacterized protein	4076	4855	-	ATG	TAA	780	259	*Pseudomonas* phage LBL3	4.6e-167
gp011	Uncharacterized protein	4941	5378	-	ATG	TAA	438	145	*Pseudomonas* phage LBL3	2.0e-91
gp012	Uncharacterized protein	5395	5982	-	ATG	TAA	588	195	*Pseudomonas* phage LBL3	4.8e-136
gp013	Uncharacterized protein	5992	6087	-	ATG	TGA	96	31	*Pseudomonas* phage NH-4	1.6e-14
gp014	Uncharacterized protein	6084	7016	-	ATG	TGA	933	310	*Pseudomonas* phage LMA2	7.4e-196
gp015	Uncharacterized protein	7120	7467	-	ATG	TGA	348	115	*Pseudomonas* phage PB1	2.6e-68
gp016	Uncharacterized protein	7716	8027	-	ATG	TAA	312	103	*Pseudomonas* phage PB1	5.8e-61
gp017	Uncharacterized protein	8033	8236	-	ATG	TAG	204	67	*Pseudomonas* phage NH-4	1.7e-37
gp018	Uncharacterized protein	8233	8556	-	ATG	TGA	324	107	*Pseudomonas* phage 14-1	2.7e-71
gp019	Uncharacterized protein	8588	8989	-	ATG	TAG	402	133	*Pseudomonas* phage 14-1	3.3e-85
gp020	Minor head protein	9170	11467	+	ATG	TGA	2298	765	*Pseudomonas* phage JG024	0.0e+00
gp021	Minor head protein	11467	12303	+	ATG	TGA	837	278	*Pseudomonas* phage SN	2.3e-182
gp022	Uncharacterized protein	12322	12528	+	ATG	TGA	207	68	*Pseudomonas* phage LBL3	1.9e-36
gp023	Uncharacterized protein	12525	12665	+	ATG	TAA	141	46	*Pseudomonas* phage LMA2	8.2e-24
gp024	Structural protein	13217	14611	+	ATG	TAA	1395	464	*Pseudomonas* phage NH-4	1.6e-304
gp025	Structural protein	14615	15250	+	ATG	TAA	636	211	*Pseudomonas* phage NH-4	2.2e-133
gp026	Structural protein	15260	16408	+	ATG	TAA	1149	382	*Pseudomonas* phage NH-4	7.3e-254
gp027	Uncharacterized protein	16510	16947	+	ATG	TAA	438	145	*Pseudomonas* phage NH-4	3.4e-91
gp028	Structural protein	16962	17429	+	ATG	TGA	468	155	*Pseudomonas* phage LMA2	3.0e-103
gp029	Structural protein	17426	17824	+	GTG	TAG	399	132	*Pseudomonas* phage SN	8.9e-86
gp030	Uncharacterized protein	17832	18383	+	ATG	TGA	552	183	*Pseudomonas* phage JG024	2.0e-119
gp031	Uncharacterized protein	18380	18961	+	ATG	TAA	582	193	*Pseudomonas* phage LMA2	6.2e-124
gp032	Structural protein	18977	20491	+	GTG	TAA	1515	504	*Pseudomonas* phage KPP12	0.0e+00

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp033	Structural protein	20550	21002	+	ATG	TGA	453	150	*Pseudomonas* phage SN	7.4e-94
gp034	Structural protein	21002	21325	+	ATG	TGA	324	107	*Pseudomonas* phage SN	6.7e-71
gp035	Structural protein	21322	21672	+	ATG	TGA	351	116	*Pseudomonas* phage SN	1.0e-70
gp036	Uncharacterized protein	21674	22105	+	ATG	TAA	432	143	*Pseudomonas* phage SN	1.0e-82
gp037	Structural protein	22115	22618	+	ATG	TGA	504	167	*Pseudomonas* phage SN	4.4e-103
gp038	Structural protein	22753	23157	+	ATG	TGA	405	134	*Pseudomonas* phage SN	2.9e-82
gp039	Structural protein	23166	23759	+	TTG	TGA	594	197	*Pseudomonas* phage SN	2.3e-121
gp040	Uncharacterized protein	23769	24197	+	ATG	TGA	429	142	*Pseudomonas* phage SN	8.7e-92
gp041	Transglycosylase	24201	26777	+	ATG	TGA	2577	858	*Pseudomonas* phage SN	0.0e+00
gp042	Structural protein	26777	27640	+	ATG	TAA	864	287	*Pseudomonas* phage SN	1.7e-189
gp043	Uncharacterized protein	27640	28173	+	ATG	TAA	534	177	*Pseudomonas* phage 14-1	8.0e-112
gp044	Baseplate protein	28229	28894	+	ATG	TAA	666	237	*Pseudomonas* phage SN	5.7e-133
gp045	Baseplate protein	28951	30204	+	ATG	TGA	1254	417	P*seudomonas* phage KPP12	3.0e-268
gp046	Structural protein	30201	31715	+	ATG	TAA	1515	504	*Pseudomonas* phage SN	0.0e+00
gp047	Tail protein	31720	34608	+	GTG	TAA	2889	962	*Pseudomonas* phage SN	0.0e+00
gp048	Uncharacterized protein	34610	35038	+	ATG	TGA	429	142	*Pseudomonas* phage SN	2.2e-93
gp049	Endolysin	35038	35700	+	ATG	TGA	663	220	*Pseudomonas* phage SN	1.5e-147
gp050	Uncharacterized protein	35724	35975	-	ATG	TAA	252	83	*Pseudomonas* phage JG024	6.7e-45
gp051	DNA ligase	36255	37166	-	ATG	TGA	912	317	*Pseudomonas* phage SN	9.6e-201
gp052	Uncharacterized protein	37221	37775	-	ATG	TGA	555	184	*Pseudomonas* phage SN	2.1e-115
gp053	Uncharacterized protein	37772	38377	-	ATG	TGA	606	201	*Pseudomonas* phage LBL3	3.4e-119
gp054	Uncharacterized protein	38431	39330	-	ATG	TAA	900	299	*Pseudomonas* phage LMA2	1.7e-194
gp055	Uncharacterized protein	30419	40051	-	TTG	TGA	633	210	*Pseudomonas* phage NH-4	3.3e-130
gp056	Helicase	40134	41693	-	ATG	TAA	1560	519	*Pseudomonas* phage LMA2	0.0e+00
gp057	DNA helicase	41690	42106	-	ATG	TGA	417	138	*Pseudomonas* phage NH-4	7.0e-88
gp058	DNA polymerase III alpha subunit	42093	45203	-	ATG	TGA	3111	1080	*Pseudomonas* phage NH-4	0.0e+00
gp059	DNA polymerase III epsilon subunit	45200	45754	-	ATG	TGA	555	184	*Pseudomonas* phage LMA2	1.8e-123
gp060	Polynucleotide kinase	45830	46846	-	ATG	TAA	1017	338	*Pseudomonas* phage KPP12	1.1e-223
gp061	Uncharacterized protein	46849	47040	-	ATG	TGA	192	63	*Pseudomonas* phage LMA2	1.7e-32
gp062	Thymidylate synthase	47042	47959	-	ATG	TAA	918	305	*Pseudomonas* phage KPP12	1.1e-205
gp063	Uncharacterized protein	47959	48165	-	ATG	TGA	207	68	*Pseudomonas* phage KPP12	5.1e-37

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp064	Uncharacterized protein	48173	48439	-	ATG	TGA	267	88	*Pseudomonas* phage PB1	3.7e-53
gp065	Uncharacterized protein	48439	48657	-	ATG	TGA	219	72	*Pseudomonas* phage KPP12	6.7e-42
gp066	Uncharacterized protein	48641	48859	-	ATG	TAA	219	72	*Pseudomonas* phage KPP12	1.3e-40
gp067	Uncharacterized protein	48859	49089	-	ATG	TAA	231	76	*Pseudomonas* phage NH-4	2.3e-46
gp068	Uncharacterized protein	49177	50178	-	ATG	TAA	1002	333	*Pseudomonas* phage LMA2	1.5e-224
gp069	Structural protein	50282	51175	-	ATG	TAA	894	297	*Pseudomonas* phage NH-4	1.8e-158
gp070	ATP-dependent exonuclease V	51334	52521	-	ATG	TGA	1188	395	*Pseudomonas* phage JG024	8.8e-271
gp071	Uncharacterized protein	52508	52930	-	ATG	TAA	423	140	*Pseudomonas* phage LMA2	1.2e-89
gp072	Uncharacterized protein	53099	53884	+	ATG	TGA	786	261	*Pseudomonas* phage 14-1	1.2e-174
gp073	Uncharacterized protein	53894	54865	+	ATG	TAA	972	323	*Pseudomonas* phage SN	3.4e-176
gp074	Uncharacterized protein	54889	55347	+	ATG	TGA	459	152	*Pseudomonas* phage LMA2	5.4e-96
gp075	Uncharacterized protein	55344	56420	+	ATG	TAG	1077	358	*Pseudomonas* phage LBL3	4.3e-254
gp076	Uncharacterized protein	56426	56611	+	ATG	TAA	186	61	*Pseudomonas* phage PB1	5.5e-32
gp077	Primase	56759	58498	+	ATG	TGA	1740	579	*Pseudomonas* phage KPP12	0.0e+00
gp078	Uncharacterized protein	59462	60031	+	ATG	TAG	570	189	*Pseudomonas* phage LMA2	4.7e-123
gp079	Uncharacterized protein	60199	60813	-	ATG	TAA	615	204	*Pseudomonas* phage LMA2	5.6e-124
gp080	Uncharacterized protein	61003	61707	-	ATG	TGA	705	234	*Pseudomonas* phage NH-4	3.4e-152
gp081	Uncharacterized protein	61718	62029	-	ATG	TGA	312	103	*Pseudomonas* phage SN	2.3e-63
gp082	Uncharacterized protein	62082	62303	-	ATG	TAA	222	73	*Pseudomonas* phage 14-1	1.9e-43
gp083	Uncharacterized protein	62313	62570	-	ATG	TAG	258	85	*Pseudomonas* phage LBL3	3.6e-43
gp084	Uncharacterized protein	62624	62848	-	ATG	TAA	225	74	*Pseudomonas* phage LMA2	9.2e-44
gp085	Uncharacterized protein	62910	63239	-	ATG	TGA	330	119	*Pseudomonas* phage LMA2	6.8e-67
gp086	Uncharacterized protein	63240	63890	-	ATG	TAA	651	216	*Pseudomonas* phage NH-4	5.0e-134
gp087	Uncharacterized protein	63922	64134	-	ATG	TAA	213	70	*Pseudomonas* phage LBL3	2.8e-43
gp088	Uncharacterized protein	64131	64331	-	ATG	TGA	201	74	*Pseudomonas* phage SN	4.9e-37
gp089	Uncharacterized protein	64328	64543	-	ATG	TGA	216	71	*Pseudomonas* phage LMA2	3.6e-41
gp090	Uncharacterized protein	64540	64740	-	GTG	TGA	201	66	*Pseudomonas* phage JG024	2.0e-38
gp091	Uncharacterized protein	64737	64988	-	ATG	TGA	252	84	*Pseudomonas* phage NH-4	3.1e-51
gp092	Uncharacterized protein	65117	65302	-	ATG	TGA	186	61	*Pseudomonas* phage LMA2	3.7e-34
gp093	Uncharacterized protein	65387	65575	-	ATG	TAA	189	62	*Pseudomonas* phage LMA2	3.3e-35
gp094	Uncharacterized protein	65578	66327	-	ATG	TGA	750	249	*Pseudomonas* phage NH-4	8.7e-148

(gp) gene product, (nt) nucleotides (aa) amino acids

Table S 4.3: The *Pseudomonas* phage KTN6 genome annotation.

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
1	Uncharacterized protein	1	303	-	ATG	TAA	303	100	*Pseudomonas* phage NH-4	9.7e-60
2	Uncharacterized protein	315	461	-	ATG	TAA	147	48	*Pseudomonas* phage LMA2	8.7e-23
3	Terminase large subunit	638	2020	+	ATG	TAA	1383	460	*Pseudomonas* phage NH-4	2.4e-314
4	Uncharacterized protein	2057	2440	-	ATG	TAA	384	127	*Pseudomonas* phage KPP12	5.7e-75
5	Uncharacterized protein	2440	2655	-	ATG	TAA	216	71	*Pseudomonas* phage SN	2.4e-41
6	Uncharacterized protein	2655	3005	-	ATG	TGA	351	116	*Pseudomonas* phage PB1	2.1e-69
7	Uncharacterized protein	3050	3481	-	ATG	TGA	432	143	*Pseudomonas* phage PB1	1.5e-76
8	Uncharacterized protein	3484	4263	-	ATG	TAA	780	259	*Pseudomonas* phage LBL3	5.3e-167
9	Uncharacterized protein	4349	4786	-	ATG	TAA	438	145	*Pseudomonas* phage LBL3	1.8e-89
10	Uncharacterized protein	4803	5390	-	ATG	TAA	588	195	*Pseudomonas* phage SN	3.7e-134
11	Uncharacterized protein	5492	6424	-	ATG	TGA	933	310	*Pseudomonas* phage LMA2	8.5e-196
12	Uncharacterized protein	6528	6875	-	ATG	TGA	348	115	*Pseudomonas* phage PB1	3.0e-68
13	Uncharacterized protein	7124	7435	-	ATG	TAA	312	103	*Pseudomonas* phage PB1	1.2e-59
14	Uncharacterized protein	7441	7644	-	ATG	TTG	204	67	*Pseudomonas* phage NH-4	2.0e-37
15	Uncharacterized protein	7641	7964	-	ATG	TGA	324	107	*Pseudomonas* phage 14-1	3.1e-71
16	Uncharacterized protein	7996	8397	-	ATG	TTG	402	133	*Pseudomonas* phage 14-1	3.8e-85
17	Minor head protein	8686	10875	+	ATG	TGA	2190	766	*Pseudomonas* phage NH-4	0.0e+00
18	Minor head protein	10875	11711	+	ATG	TGA	837	278	*Pseudomonas* phage SN	3.5e-182
19	Uncharacterized protein	11730	11936	+	ATG	TGA	207	68	*Pseudomonas* phage 14-1	2.2e-36
20	Uncharacterized protein	11933	12073	+	ATG	TAA	141	46	*Pseudomonas* phage NH-4	9.5e-24
21	Structural protein	12586	14019	+	ATG	TAA	1395	464	*Pseudomonas* phage NH-4	1.8e-304
22	Structural protein	14023	14658	+	ATG	TAA	636	211	*Pseudomonas* phage LMA2	2.5e-133
23	Major structural protein	14668	15816	+	ATG	TAA	1149	382	*Pseudomonas* phage NH-4	1.1e-253
24	Uncharacterized protein	15918	16355	+	ATG	TAA	438	145	*Pseudomonas* phage 14-1	5.4e-91
25	Structural protein	16370	16837	+	ATG	TGA	468	155	*Pseudomonas* phage KPP12	4.9e-103
26	Structural protein	16834	17232	+	GTG	TGA	399	132	*Pseudomonas* phage KPP12	8.6e-85
27	Structural protein	17240	17791	+	ATG	TAA	552	183	*Pseudomonas* phage SN	1.0e-117
28	Uncharacterized protein	17788	18369	+	GTG	TAA	582	193	*Pseudomonas* phage SN	7.1e-124
30	Structural protein	19958	20410	+	ATG	TGA	453	150	*Pseudomonas* phage KPP12	5.9e-95
31	Structural protein	20410	20733	+	ATG	TGA	324	107	*Pseudomonas* phage NH-4	7.7e-71
32	Structural protein	20730	21080	+	ATG	TGA	351	116	*Pseudomonas* phage NH-4	1.2e-70
33	Uncharacterized protein	21082	21513	+	ATG	TAA	432	143	*Pseudomonas* phage SN	2.2e-83

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
34	Structural protein	21523	22026	+	ATG	TGA	504	167	*Pseudomonas* phage SN	5.1e-103
35	Structural protein	22161	22565	+	ATG	TGA	540	180	*Pseudomonas* phage LMA2	8.8e-112
36	Structural protein	22574	23167	+	TTG	TGA	594	197	*Pseudomonas* phage KPP12	8.6e-121
37	Uncharacterized protein	23177	23605	+	ATG	TGA	429	142	*Pseudomonas* phage LBL3	2.6e-91
38	Transglycosydase	23621	26179	+	TTG	TGA	2571	857	*Pseudomonas* phage JG024	0.0e+00
39	Structural protein	26179	27042	+	ATG	TAA	864	287	*Pseudomonas* phage LMA2	7.3e-190
40	Uncharacterized protein	27042	27575	+	ATG	TAA	534	177	*Pseudomonas* phage KPP12	1.1e-110
41	Baseplate	27733	28296	+	ATG	TAA	666	222	*Pseudomonas* phage KPP12	2.4e-143
42	Baseplate	28353	29606	+	ATG	TGA	1254	417	*Pseudomonas* phage KPP12	3.4e-268
43	Structural protein	29603	31117	+	ATG	TAA	1515	504	*Pseudomonas* phage KPP12	0.0e+00
44	Tail fibers protein	31122	34016	+	GTG	TAA	2895	964	*Pseudomonas* phage SPM-1	0.0e+00
45	Tail fiber component	34018	34446	+	ATG	TGA	429	142	*Pseudomonas* phage KPP12	5.8e-90
46	Endolysin	34446	35108	+	ATG	TGA	663	220	*Pseudomonas* phage KPP12	1.5e-147
47	Uncharacterized protein	35133	35384	-	ATG	TAA	252	83	*Pseudomonas* phage JG024	7.7e-45
48	DNA ligase	35664	36566	-	ATG	TGA	912	304	*Pseudomonas* phage LMA2	8.4e-212
49	Uncharacterized protein	36630	37184	-	ATG	TGA	555	184	*Pseudomonas* phage NH-4	1.1e-118
50	Uncharacterized protein	37181	37786	-	ATG	TGA	606	201	*Pseudomonas* phage PB1	3.9e-119
51	Uncharacterized protein	37840	38739	-	ATG	TAA	900	299	*Pseudomonas* phage 14-1	1.9e-194
52	Uncharacterized protein	38828	39448	-	ATG	TGA	621	206	*Pseudomonas* phage KPP12	2.0e-128
53	Helicase	39543	41102	-	ATG	TAA	1560	519	*Pseudomonas* phage LMA2	0.0e+00
54	DNA helicase	41099	41509	-	ATG	TGA	411	136	*Pseudomonas* phage SN	1.3e-88
55	DNA Polymerase III alpha subunit	41502	44609	-	ATG	TGA	3111	1037	*Pseudomonas* phage SN	0.0e+00
56	DNA polymerase III epsilon subunit	44609	45163	-	ATG	TGA	555	184	*Pseudomonas* phage KPP12	7.7e-124
57	Polynucleotide kinase	45239	46258	-	ATG	TAA	1020	339	*Pseudomonas* phage KPP12	7.9e-222
58	Uncharacterized protein	46261	46452	-	ATG	TGA	192	63	*Pseudomonas* phage SN	2.6e-33
59	Thymidylate synthase	46454	47371	-	ATG	TAA	918	305	*Pseudomonas* phage LMA2	1.6e-206
60	Uncharacterized protein	47371	47577	-	ATG	TGA	207	68	*Pseudomonas* phage LMA2	1.9e-38
61	Uncharacterized protein	47585	47842	-	ATG	TGA	258	86	*Pseudomonas* phage NH-4	5.8e-46
62	Tail assembly protein	47842	48060	-	ATG	TAA	219	72	*Pseudomonas* phage SPM-1	1.4e-37
63	Uncharacterized protein	48044	48262	-	ATG	TAA	219	72	*Pseudomonas* phage NH-4	1.4e-40
64	Uncharacterized protein	48262	48492	-	ATG	TAA	231	76	*Pseudomonas* phage NH-4	3.4e-46

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
65	Uncharacterized protein	48580	49581	-	ATG	TAA	1002	333	*Pseudomonas* phage KPP12	3.6e-224
66	Structural protein	49686	50579	-	ATG	TAA	894	297	*Pseudomonas* phage NH-4	4.0e-154
67	ATP-dependent exonuclease V	50740	51927	-	ATG	TGA	1188	395	*Pseudomonas* phage JG024	7.2e-270
68	Uncharacterized protein	51914	52336	-	ATG	TAA	423	140	*Pseudomonas* phage LMA2	9.1e-92
69	Uncharacterized protein	52505	53290	+	ATG	TGA	786	261	*Pseudomonas* phage 14-1	9.6e-175
70	Uncharacterized protein	53292	54485	+	GTG	TAA	1185	395	*Pseudomonas* phage LMA2	2.2e-263
71	Uncharacterized protein	54509	54967	+	ATG	TGA	459	152	*Pseudomonas* phage SPM-1	5.5e-97
72	Uncharacterized protein	54964	56040	+	ATG	TAG	1077	358	*Pseudomonas* phage KPP12	5.3e-254
73	Uncharacterized protein	56046	56231	+	ATG	TAG	186	61	*Pseudomonas* phage LBL3	2.6e-30
74	Primase	56379	58118	+	ATG	TGA	1740	579	*Pseudomonas* phage NH-4	0.0e+00
75	Uncharacterized protein	59099	59668	+	ATG	TAG	570	189	*Pseudomonas* phage LMA2	1.1e-123
76	Uncharacterized protein	59835	60449	-	ATG	TAA	615	204	*Pseudomonas* phage LMA2	6.4e-124
77	Uncharacterized protein	60680	61339	-	ATG	TAA	660	219	*Pseudomonas* phage NH-4	4.8e-136
78	Uncharacterized protein	61350	61661	-	ATG	TGA	312	103	*Pseudomonas* phage SN	5.4e-65
79	Uncharacterized protein	61714	61935	-	ATG	TAA	222	73	*Pseudomonas* phage LBL3	5.4e-43
80	Uncharacterized protein	61945	62202	-	ATG	TTG	258	85	*Pseudomonas* phage NH-4	2.6e-47
81	Uncharacterized protein	62256	62480	-	ATG	TAA	225	74	*Pseudomonas* phage LMA2	1.1e-43
82	Uncharacterized protein	62542	62865	-	ATG	TGA	324	108	*Pseudomonas* phage LMA2	2.2e-66
83	Uncharacterized protein	62868	63515	-	ATG	TGA	648	215	*Pseudomonas* phage PB1	3.7e-145
84	Uncharacterized protein	63547	63759	-	ATG	TAA	213	70	*Pseudomonas* phage KPP12	1.7e-43
85	Uncharacterized protein	63756	63944	-	ATG	TGA	189	62	*Pseudomonas* phage LBL3	3.0e-38
86	Uncharacterized protein	63941	64156	-	ATG	TGA	216	71	*Pseudomonas* phage LBL3	1.8e-41
87	Uncharacterized protein	64153	64353	-	GTG	TGA	201	66	*Pseudomonas* phage LBL3	2.8e-38
88	Uncharacterized protein	64350	64601	-	ATG	TGA	252	83	*Pseudomonas* phage NH-4	3.6e-51
89	Uncharacterized protein	64730	64915	-	ATG	TGA	186	61	*Pseudomonas* phage LMA2	4.2e-34
90	Uncharacterized protein	65000	65188	-	ATG	TAA	189	62	*Pseudomonas* phage LMA2	3.8e-35
91	Tail length tape-measure protein	65191	65976	-	ATG	TGA	750	250	*Pseudomonas* phage SPM-1	2.9e-65

(gp) gene product, (nt) nucleotides (aa) amino acids

Table S 5.1: The KTN4 genome annotation.

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp001	Virion associated protein	1	390	+	ATG	TAA	389	129	PHIKZ001	*Pseudomonas* phage φKZ	9.6e-84
gp002	Uncharacterized protein	354	938	+	ATG	TAA	584	194	PHIKZ002	*Pseudomonas* phage φKZ	1.5e-127
gp003	Uncharacterized protein	913	1419	+	ATG	TAA	506	168	PHIKZ003	*Pseudomonas* phage φKZ	8.4e-107
gp004	Uncharacterized protein	1419	1940	+	ATG	TAA	521	173	PHIKZ004	*Pseudomonas* phage φKZ	2.1e-114
gp005	Uncharacterized protein	2057	2461	+	ATG	TAA	404	134	PHIKZ005	*Pseudomonas* phage φKZ	6.2e-86
gp006	Uncharacterized protein	2556	3071	+	ATG	TAA	515	171	PHIKZ006	*Pseudomonas* phage φKZ	2.6e-120
gp007	Uncharacterized protein	3089	3325	+	ATG	TAA	236	78	PHIKZ007	*Pseudomonas* phage φKZ	1.0e-47
gp008	Uncharacterized protein	3328	3540	+	ATG	TAA	212	70	PHIKZ008	*Pseudomonas* phage φKZ	7.6e-41
gp009	Uncharacterized protein	3642	3974	+	ATG	TGA	332	110	PHIKZ009	*Pseudomonas* phage φKZ	5.2e-36
gp010	Uncharacterized protein	4056	4316	+	ATG	TAA	260	86	PHIKZ010	*Pseudomonas* phage φKZ	8.1e-53
gp011	Uncharacterized protein	4407	4556	+	ATG	TAG	149	49	PHIKZ010.1	*Pseudomonas* phage φKZ	1.3e-25
gp012	Uncharacterized protein	4567	4908	+	ATG	TAA	341	113	PHIKZ011	*Pseudomonas* phage φKZ	2.9e-70
gp013	Uncharacterized protein	4914	5312	+	ATG	TGA	398	132	PHIKZ012	*Pseudomonas* phage φKZ	3.7e-76
gp014	Uncharacterized protein	5309	5482	+	GTG	TAA	173	57	Glycophorin binding protein family, Gbph	*Plasmodium falciparum* 3D7	1.9e-06
gp015	Uncharacterized protein	5486	5893	+	ATG	TGA	407	135	PHIKZ013	*Pseudomonas* phage φKZ	1.7e-83
gp016	Uncharacterized protein	6043	7155	+	ATG	TAG	1112	370	PHIKZ014	*Pseudomonas* phage φKZ	3.6e-252
gp017	Uncharacterized protein	7260	8261	+	ATG	TGA	1001	333	PHIKZ015	*Pseudomonas* phage φKZ	2.5e-211
gp018	Uncharacterized protein	8274	8753	+	ATG	TAA	479	159	PHIKZ016	*Pseudomonas* phage φKZ	2.6e-99
gp019	Uncharacterized protein	8810	8953	+	ATG	TAA	143	47	PHIKZ016.1	*Pseudomonas* phage φKZ	2.0e-25
gp020	Uncharacterized protein	8982	9374	+	ATG	TAA	392	130	PHIKZ017	*Pseudomonas* phage φKZ	3.7e-78
gp021	Uncharacterized protein	9437	9766	+	ATG	TAA	329	109	PHIKZ018	*Pseudomonas* phage φKZ	2.1e-69
gp022	Uncharacterized protein	9796	10005	+	ATG	TAA	209	69	PHIKZ019	*Pseudomonas* phage φKZ	1.7e-40
gp023	Uncharacterized protein	10063	10350	+	ATG	TAA	287	95	Uncharacterized protein	*Pseudomonas* phage PA7	7.1e-61
gp024	Uncharacterized protein	10416	10706	+	GTG	TAG	290	96	Uncharacterized protein	*Pseudomonas* phage 201phi2-1	5.3e-09
gp025	Uncharacterized protein	10718	11056	+	ATG	TAG	338	112	Uncharacterized protein	*Pseudomonas* phage PA7	1.8e-70
gp026	Uncharacterized protein	11046	11186	+	ATG	TGA	140	46			
gp027	Uncharacterized protein	11186	11443	+	ATG	TAA	257	85	PHIKZ020	*Pseudomonas* phage φKZ	7.5e-53
gp028	Uncharacterized protein	11590	11925	+	ATG	TGA	335	111	PHIKZ021	*Pseudomonas* phage φKZ	1.1e-68
gp029	Uncharacterized protein	11933	12487	+	ATG	TGA	554	184	PHIKZ022	*Pseudomonas* phage φKZ	4.7e-114
gp030	Uncharacterized protein	12484	12693	+	ATG	TGA	209	69	PHIKZ019	*Pseudomonas* phage φKZ	2.9e-23

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp031	Uncharacterized protein	12690	12932	+	ATG	TAA	242	80			
gp032	Uncharacterized protein	12932	13198	+	ATG	TAA	266	88			
gp033	Uncharacterized protein	13204	13641	+	ATG	TAA	437	145	PHIKZ023	*Pseudomonas* phage φKZ	3.0e-81
gp034	Uncharacterized protein	13895	14161	+	ATG	TAA	266	88	Uncharacterized protein	*Pseudomonas* phage PA7	5.3e-47
gp035	Terminase large subunit	14215	16368	-	ATG	TAA	2153	717	PHIKZ025	*Pseudomonas* phage φKZ	0.0e+00
gp036	Structural head protein	16508	18166	-	ATG	TAA	1658	552	PHIKZ026	*Pseudomonas* phage φKZ	0.0e+00
gp037	Structural head protein	18163	20859	-	ATG	TGA	2696	898	PHIKZ027	*Pseudomonas* phage φKZ	0.0e+00
gp038	Structural head protein	20876	21838	-	ATG	TAA	962	320	PHIKZ028	*Pseudomonas* phage φKZ	3.1e-215
gp039	Tail sheath protein	21967	24054	+	ATG	TGA	2087	695	PHIKZ029	*Pseudomonas* phage φKZ	0.0e+00
gp040	Structural protein	24086	24967	+	ATG	TAA	881	293	PHIKZ030	*Pseudomonas* phage φKZ	2.7e-201
gp041	Uncharacterized protein	24967	25149	+	ATG	TAA	182	60	PHIKZ030.1	*Pseudomonas* phage φKZ	3.7e-34
gp042	Uncharacterized protein	25277	26473	+	ATG	TGA	1196	398	PHIKZ031	*Pseudomonas* phage φKZ	1.4e-271
gp043	Uncharacterized protein	26470	26826	+	ATG	TAA	356	118	PHIKZ031.1	*Pseudomonas* phage φKZ	4.4e-74
gp044	Structural head protein	26873	28111	-	ATG	TAA	1238	412	PHIKZ032	*Pseudomonas* phage φKZ	7.3e-275
gp045	Uncharacterized protein	28202	29071	+	ATG	TAG	869	289	PHIKZ033	*Pseudomonas* phage φKZ	4.9e-175
gp046	Uncharacterized protein	29079	29282	+	ATG	TAA	203	67	PHIKZ033.1	*Pseudomonas* phage φKZ	1.3e-39
gp047	Structural protein	29378	30148	+	ATG	TAA	770	256	PHIKZ034	*Pseudomonas* phage φKZ	3.5e-174
gp048	Structural head protein	30158	30973	+	ATG	TGA	815	271	PHIKZ035	*Pseudomonas* phage φKZ	2.5e-166
gp049	Uncharacterized protein	31077	31889	+	ATG	TAA	812	270	PHIKZ036	*Pseudomonas* phage φKZ	3.3e-180
gp050	Uncharacterized protein	31926	32747	+	ATG	TGA	821	273	PHIKZ037	*Pseudomonas* phage φKZ	6.3e-172
gp051	Uncharacterized protein	32987	33499	+	ATG	TAA	512	170	PHIKZ038	*Pseudomonas* phage φKZ	9.7e-106
gp052	Chain A, monomeric subunit of Tubz	33553	34533	+	ATG	TAA	980	326	PHIKZ039	*Pseudomonas* phage φKZ	4.5e-215
gp053	Uncharacterized protein	34588	35070	+	ATG	TAA	482	160	PHIKZ040	*Pseudomonas* phage φKZ	5.1e-104
gp054	Uncharacterized protein	35110	35583	+	ATG	TGA	473	157	PHIKZ041	*Pseudomonas* phage φKZ	3.0e-104
gp055	Uncharacterized protein	35595	35912	+	ATG	TAA	317	105	PHIKZ041.1	*Pseudomonas* phage φKZ	2.5e-63
gp056	Uncharacterized protein	35891	36178	+	ATG	TAA	287	95	PHIKZ041.2	*Pseudomonas* phage φKZ	6.9e-55
gp057	Uncharacterized protein	36246	36473	+	ATG	TGA	227	101	PHIKZ041.3	*Pseudomonas* phage φKZ	1.5e-45
gp058	Virion associated protein	36578	37378	-	ATG	TAA	800	266	PHIKZ042	*Pseudomonas* phage φKZ	1.2e-176
gp059	Uncharacterized protein	37515	37943	+	ATG	TAA	428	142			
gp060	Uncharacterized protein	37957	39030	+	ATG	TAA	1073	357	PHIKZ043	*Pseudomonas* phage φKZ	4.0e-232
gp061	Uncharacterized protein	39080	39469	+	ATG	TAA	389	129	PHIKZ044	*Pseudomonas* phage φKZ	4.2e-73
gp062	Uncharacterized protein	39393	39845	+	ATG	TAA	452	150	PHIKZ044.1	*Pseudomonas* phage φKZ	1.4e-53

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp063	Uncharacterized protein	39718	39972	+	ATG	TAG	254	84	PHIKZ045	*Pseudomonas* phage φKZ	3.4e-47
gp064	Uncharacterized protein	40118	40552	+	ATG	TGA	434	144	PHIKZ046	*Pseudomonas* phage φKZ	3.3e-94
gp065	Uncharacterized protein	40549	41085	+	ATG	TAA	536	178	PHIKZ047	*Pseudomonas* phage φKZ	5.5e-116
gp066	Uncharacterized protein	41158	41541	+	ATG	TAA	383	127	PHIKZ048	*Pseudomonas* phage φKZ	3.8e-82
gp067	Virion associated protein	41576	41992	+	ATG	TAA	416	138	PHIKZ049	*Pseudomonas* phage φKZ	3.3e-86
gp068	Uncharacterized protein	42193	44388	+	ATG	TAA	2195	731	PHIKZ050	*Pseudomonas* phage φKZ	0.0e+00
gp069	Uncharacterized protein	44487	44663	+	ATG	TAA	176	58	PHIKZ050.1	*Pseudomonas* phage φKZ	9.2e-33
gp070	Uncharacterized protein	44650	45369	+	ATG	TAA	719	239	PHIKZ051	*Pseudomonas* phage φKZ	3.3e-154
gp071	Structural head protein	45413	46504	-	ATG	TGA	1091	363	PHIKZ052	*Pseudomonas* phage φKZ	8.8e-235
gp072	Uncharacterized protein	46586	46951	+	ATG	TGA	365	121	PHIKZ053	*Pseudomonas* phage φKZ	8.4e-75
gp073	Virion associated protein	47263	49143	+	ATG	TAA	1880	626	PHIKZ054	*Pseudomonas* phage φKZ	0.0e+00
gp074	Non-virion DNA-dependent RNA polymerase subunit	49538	50677	+	ATG	TAA	1139	379	PHIKZ055	*Pseudomonas* phage φKZ	2.6e-259
gp075	Uncharacterized protein	50740	51561	+	ATG	TAA	821	273			
gp076	Uncharacterized protein	52060	53064	+	ATG	TAA	1004	334	PHIKZ056	*Pseudomonas* phage φKZ	3.7e-226
gp077	Uncharacterized protein	53435	53620	+	ATG	TAA	185	61	PHIKZ056.1	*Pseudomonas* phage φKZ	2.0e-36
gp078	Uncharacterized protein	53669	54052	+	ATG	TAA	383	127	PHIKZ057	*Pseudomonas* phage φKZ	4.2e-84
gp079	Uncharacterized protein	54065	54490	+	ATG	TGA	425	141	PHIKZ058	*Pseudomonas* phage φKZ	1.4e-93
gp080	Uncharacterized protein	54613	54795	+	ATG	TAA	182	60	PHIKZ058.1	*Pseudomonas* phage φKZ	3.7e-29
gp081	Uncharacterized protein	54865	55476	+	ATG	TAA	611	203	PHIKZ059	*Pseudomonas* phage φKZ	2.5e-133
gp082	Uncharacterized protein	55528	55863	+	ATG	TAA	335	111	PHIKZ060	*Pseudomonas* phage φKZ	1.6e-70
gp083	Uncharacterized protein	55850	56302	+	ATG	TAG	452	150	PHIKZ061	*Pseudomonas* phage φKZ	1.5e-87
gp084	Virion associated protein	56408	56818	+	ATG	TAA	410	136	PHIKZ062	*Pseudomonas* phage φKZ	5.3e-83
gp085	Virion associated protein	56819	57376	+	ATG	TAG	557	185	PHIKZ063	*Pseudomonas* phage φKZ	6.5e-111
gp086	Virion associated protein	57396	57959	+	ATG	TGA	563	187	PHIKZ064	*Pseudomonas* phage φKZ	4.9e-117
gp087	Nuclease SbcCD, D subunit	57949	59088	+	ATG	TAA	1139	379	PHIKZ065	*Pseudomonas* phage φKZ	1.6e-258
gp088	Uncharacterized protein	59101	59931	+	ATG	TAG	830	276	PHIKZ066	*Pseudomonas* phage φKZ	2.8e-183
gp089	Virion associated protein	60018	60758	+	ATG	TAA	740	246	PHIKZ067	*Pseudomonas* phage φKZ	6.9e-168
gp090	Non-virion DNA-dependent RNA polymerase subunit	60884	62449	+	ATG	TAA	1565	521	PHIKZ068	*Pseudomonas* phage φKZ	0.0e+00

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp091	Uncharacterized protein	62466	64139	+	ATG	TAG	1673	557	PHIKZ069	*Pseudomonas* phage φKZ	0.0e+00
gp092	Uncharacterized protein	64186	65487	+	ATG	TAA	1301	459	PHIKZ070	*Pseudomonas* phage φKZ	1.4e-285
gp093	Virion associated protein	65529	65891	-	ATG	TAA	362	120	Uncharacterized protein	*Pseudomonas* phage PA7	1.3e-78
gp094	Uncharacterized protein	65956	66075	+	ATG	TGA	119	36	Uncharacterized protein	*Pseudomonas* phage PA7	2.00E-18
gp095	Non-virion DNA-dependent RNA polymerase subunit	66068	67252	+	ATG	TAA	1184	399	PHIKZ071	*Pseudomonas* phage φKZ	1.1e-270
gp096	Non-virion DNA-dependent RNA polymerase subunit	67575	68477	+	ATG	TGA	902	300	PHIKZ073	*Pseudomonas* phage φKZ	1.8e-201
gp097	Non-virion DNA-dependent RNA polymerase subunit	68508	70541	+	ATG	TAG	2033	677	PHIKZ074	*Pseudomonas* phage φKZ	0.0e+00
gp098	Putative RAD2/SF2 helicase	70600	72138	+	ATG	TAG	1538	512	PHIKZ075	*Pseudomonas* phage φKZ	0.0e+00
gp099	Uncharacterized protein	72321	72578	+	ATG	TGA	257	85	PHIKZ076	*Pseudomonas* phage φKZ	8.9e-47
gp100	Uncharacterized protein	72591	72782	+	ATG	TAG	191	63	PHIKZ076.1	*Pseudomonas* phage φKZ	7.9e-31
gp101	Uncharacterized protein	73128	73271	+	ATG	TAA	143	47	PHIKZ076.3	*Pseudomonas* phage φKZ	9.7e-22
gp102	Uncharacterized protein	73322	73564	-	ATG	TAA	242	80	PHIKZ076.4	*Pseudomonas* phage φKZ	2.3e-48
gp103	Uncharacterized protein	73783	75384	-	ATG	TAG	1601	533	PHIKZ077	*Pseudomonas* phage φKZ	0.0e+00
gp104	Virion associated protein	75396	75890	-	ATG	TAA	494	164	PHIKZ078	*Pseudomonas* phage φKZ	4.3e-100
gp105	Structural head protein	76062	76919	+	ATG	TGA	857	285	PHIKZ079	*Pseudomonas* phage φKZ	6.1e-172
gp106	β/β′-like virion-associated proteins	76906	78255	+	ATG	TAA	1349	449	PHIKZ080	*Pseudomonas* phage φKZ	1.7e-311
gp107	Structural protein	78318	79739	+	ATG	TAA	1421	473	PHIKZ081	*Pseudomonas* phage φKZ	6.2e-311
gp108	Uncharacterized protein	79841	81757	+	ATG	TAA	1916	638	PHIKZ082	*Pseudomonas* phage φKZ	0.0e+00
gp109	Uncharacterized protein	82009	82125	+	ATG	TAA	116	38	PHIKZ082.1	*Pseudomonas* phage φKZ	2.7e-18
gp110	Structural head protein	82229	83743	-	ATG	TAA	1514	504	PHIKZ083	*Pseudomonas* phage φKZ	0.0e+00
gp111	Structural head protein	83845	85086	+	ATG	TAG	1241	413	PHIKZ084	*Pseudomonas* phage φKZ	1.0e-271
gp112	Structural head protein	85098	85547	+	ATG	TGA	449	149	PHIKZ085	*Pseudomonas* phage φKZ	8.7e-89
gp113	Structural head protein	85558	86826	+	ATG	TAA	1268	422	PHIKZ086	*Pseudomonas* phage φKZ	6.6e-278
gp114	Structural protein	86868	89783	-	ATG	TAA	2915	971	PHIKZ087	*Pseudomonas* phage φKZ	0.0e+00
gp115	Structural protein	89787	90836	-	ATG	TAA	1049	349	PHIKZ088	*Pseudomonas* phage φKZ	3.5e-239

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp116	Stuctural head protein	90876	92039	+	ATG	TAA	1163	387	PHIKZ089	*Pseudomonas* phage φKZ	2.1e-252
gp117	Structural head protein	92050	92979	+	ATG	TAA	929	309	PHIKZ090	*Pseudomonas* phage φKZ	1.6e-195
gp118	Structural head protein	92994	93524	+	ATG	TGA	530	176	PHIKZ091	*Pseudomonas* phage φKZ	4.4e-117
gp119	Structural head protein	93534	94856	+	ATG	TAA	1322	440	PHIKZ092	*Pseudomonas* phage φKZ	1.7e-288
gp120	Structural head protein	94862	96169	+	ATG	TAA	1307	435	PHIKZ093	*Pseudomonas* phage φKZ	1.3e-270
gp121	Structural head protein	96239	97756	+	ATG	TAA	1517	505	PHIKZ094	*Pseudomonas* phage φKZ	0.0e+00
gp122	Structural head protein	97844	99487	+	ATG	TAA	1643	547	PHIKZ095	*Pseudomonas* phage φKZ	0.0e+00
gp123	Structural head protein	99549	100664	+	ATG	TAA	1115	371	PHIKZ096	*Pseudomonas* phage φKZ	5.0e-240
gp124	Structural head protein	100764	103214	+	ATG	TAA	2450	816	PHIKZ097	*Pseudomonas* phage φKZ	0.0e+00
gp125	Virion associated protein	103298	104926	+	ATG	TAA	1628	542	PHIKZ098	*Pseudomonas* phage φKZ	0.0e+00
gp126	Structural head protein	104926	106335	+	ATG	TAA	1409	469	PHIKZ099	*Pseudomonas* phage φKZ	0.0e+00
gp127	Virion associated protein	106340	106921	+	ATG	TGA	581	193	PHIKZ100	*Pseudomonas* phage φKZ	8.4e-128
gp128	Structural protein	106923	108305	+	ATG	TAA	1382	460	PHIKZ101	*Pseudomonas* phage φKZ	9.7e-311
gp129	Uncharacterized protein	108330	108752	+	ATG	TAA	422	140	PHIKZ102	*Pseudomonas* phage φKZ	1.4e-93
gp130	Uncharacterized protein	108792	108917	+	ATG	TAG	125	41	PHIKZ102.1	*Pseudomonas* phage φKZ	3.9e-21
gp131	Uncharacterized protein	108926	109318	+	ATG	TAA	392	130	PHIKZ103	*Pseudomonas* phage φKZ	1.5e-83
gp132	Virion associated protein	109399	109887	+	ATG	TAA	488	162	PHIKZ104	*Pseudomonas* phage φKZ	1.9e-109
gp133	Uncharacterized protein	110010	110348	+	ATG	TAA	338	112	PHIKZ106	*Pseudomonas* phage φKZ	6.6e-65
gp134	Uncharacterized protein	110457	110618	+	ATG	TAA	161	53	PHIKZ106.1	*Pseudomonas* phage φKZ	1.7e-26
gp135	Uncharacterized protein	110664	110981	+	ATG	TAA	317	105	PHIKZ107	*Pseudomonas* phage φKZ	1.1e-58
gp136	Uncharacterized protein	110971	111411	+	ATG	TAA	440	146	PHIKZ107.1	*Pseudomonas* phage φKZ	4.6e-57
gp137	Uncharacterized protein	111499	111855	+	ATG	TAA	356	118	PHIKZ108	*Pseudomonas* phage φKZ	4.2e-73
gp138	Uncharacterized protein	111866	112288	+	ATG	TGA	422	140	PHIKZ109	*Pseudomonas* phage φKZ	3.1e-89
gp139	Uncharacterized protein	112360	112611	+	ATG	TAG	251	83	PHIKZ109.1	*Pseudomonas* phage φKZ	4.8e-51
gp140	Uncharacterized protein	112647	113030	+	ATG	TGA	383	127	PHIKZ110	*Pseudomonas* phage φKZ	1.7e-82
gp141	Uncharacterized protein	113066	113563	+	ATG	TAA	497	165	PHIKZ111	*Pseudomonas* phage φKZ	5.8e-110
gp142	Uncharacterized protein	113605	113880	+	ATG	TAG	275	91	PHIKZ112	*Pseudomonas* phage φKZ	8.3e-59
gp143	Uncharacterized protein	113927	114190	+	ATG	TAA	263	87	PHIKZ113	*Pseudomonas* phage φKZ	2.4e-56
gp144	Uncharacterized protein	114198	114482	+	ATG	TAA	284	94	PHIKZ114	*Pseudomonas* phage φKZ	8.0e-56
gp145	Uncharacterized protein	114546	114773	+	ATG	TGA	227	75	PHIKZ115	*Pseudomonas* phage φKZ	2.5e-47
gp146	Uncharacterized protein	114776	114931	+	ATG	TGA	155	51	PHIKZ115.1	*Pseudomonas* phage φKZ	6.0e-27
gp147	Uncharacterized protein	114928	115140	+	GTG	TAA	212	70	PHIKZ115.2	*Pseudomonas* phage φKZ	1.1e-42
gp148	Uncharacterized protein	115180	115323	+	ATG	TAA	143	47	PHIKZ115.3	*Pseudomonas* phage φKZ	7.7e-28

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp149	Uncharacterized protein	115325	115480	+	ATG	TGA	155	51	PHIKZ116	*Pseudomonas* phage φKZ	4.2e-29
gp150	Uncharacterized protein	115765	116070	+	ATG	TGA	305	101			
gp151	Uncharacterized protein	116714	116950	-	ATG	TAA	236	78	PHIKZ117.3	*Pseudomonas* phage φKZ	3.2e-45
gp152	Uncharacterized protein	116952	117320	-	ATG	TAA	368	122	PHIKZ117.4	*Pseudomonas* phage φKZ	3.3e-79
gp153	DnaB helicase	117375	118922	-	ATG	TAA	1547	515	PHIKZ118	*Pseudomonas* phage φKZ	0.0e+00
gp154	Structural head protein	119062	119595	+	ATG	TAG	533	177	PHIKZ119	*Pseudomonas* phage φKZ	8.7e-113
gp155	Major head protein	119668	121911	+	ATG	TGA	2243	747	PHIKZ120	*Pseudomonas* phage φKZ	0.0e+00
gp156	Uncharacterized protein	122108	122560	+	ATG	TAA	452	150	PHIKZ121	*Pseudomonas* phage φKZ	4.7e-95
gp157	Uncharacterized protein	122517	123599	+	ATG	TAG	1082	360	PHIKZ122	*Pseudomonas* phage φKZ	1.4e-244
gp158	non-virion DNA-dependent RNA polymerase subunit	123609	125240	+	ATG	TAA	1631	543	PHIKZ123	*Pseudomonas* phage φKZ	0.0e+00
gp159	Virion associated protein	125276	127441	+	ATG	TAA	2165	721	PHIKZ124	*Pseudomonas* phage φKZ	0.0e+00
gp160	Uncharacterized protein	127522	128154	+	ATG	TAA	632	210	PHIKZ125	*Pseudomonas* phage φKZ	7.3e-142
gp161	Structural protein	128219	128662	-	ATG	TAA	443	147	PHIKZ126	*Pseudomonas* phage φKZ	5.7e-94
gp162	Structural protein	128671	129543	-	ATG	TGA	872	290	PHIKZ127	*Pseudomonas* phage φKZ	6.1e-194
gp163	Structural head protein	129554	131728	-	ATG	TAA	2174	724	PHIKZ128	*Pseudomonas* phage φKZ	0.0e+00
gp164	Structural protein	131764	134445	+	ATG	TGA	2681	893	PHIKZ129	*Pseudomonas* phage φKZ	0.0e+00
gp165	Structural protein	134495	135778	-	ATG	TAA	1283	427	PHIKZ130	*Pseudomonas* phage φKZ	4.3e-296
gp166	Tail tip protein	135876	138191	+	ATG	TAA	2315	771	PHIKZ131	*Pseudomonas* phage φKZ	0.0e+00
gp167	Structural protein	138204	138536	+	ATG	TAA	332	110	PHIKZ132	*Pseudomonas* phage φKZ	3.6e-68
gp168	Structural protein	138611	139999	+	ATG	TAA	1388	462	PHIKZ133	*Pseudomonas* phage φKZ	0.0e+00
gp169	Structural protein	140011	141384	+	ATG	TAA	1373	457	PHIKZ134	*Pseudomonas* phage φKZ	0.0e+00
gp170	Structural protein	141384	142775	+	ATG	TAA	1391	463	PHIKZ135	*Pseudomonas* phage φKZ	0.0e+00
gp171	Virion associated protein	142775	143377	+	ATG	TGA	602	200	PHIKZ136	*Pseudomonas* phage φKZ	1.8e-127
gp172	Uncharacterized protein	143364	143810	+	GTG	TAG	446	148	PHIKZ137	*Pseudomonas* phage φKZ	1.8e-99
gp173	Virion associated protein	143819	144307	+	ATG	TAA	488	162	PHIKZ138	*Pseudomonas* phage φKZ	9.5e-108
gp174	Structural protein	144403	145299	+	ATG	TAA	896	298	PHIKZ139	*Pseudomonas* phage φKZ	1.4e-197
gp175	Uncharacterized protein	145302	145865	+	ATG	TAA	563	187	PHIKZ140	*Pseudomonas* phage φKZ	3.4e-122
gp176	Structural protein	145881	146387	+	ATG	TAG	506	168	PHIKZ141	*Pseudomonas* phage φKZ	2.5e-105
gp177	Uncharacterized protein	146407	146634	+	GTG	TGA	227	75	PHIKZ141.1	*Pseudomonas* phage φKZ	2.0e-45
gp178	Uncharacterized protein	146609	147004	+	ATG	TAA	395	131	PHIKZ142	*Pseudomonas* phage φKZ	2.4e-74
gp179	Structural protein	147086	147619	+	ATG	TAA	533	177	PHIKZ143	*Pseudomonas* phage φKZ	1.0e-113

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp180	Endolysin	147681	148463	+	ATG	TAA	782	260	PHIKZ144	*Pseudomonas* phage φKZ	1.5e-175
gp181	Putative tail sheath protein	148669	151038	+	ATG	TAA	2369	789	PHIKZ145	*Pseudomonas* phage φKZ	0.0e+00
gp182	Putative tail fiber protein	151038	154319	+	ATG	TAA	3281	1093	PHIKZ146	*Pseudomonas* phage φKZ	0.0e+00
gp183	Uncharacterized protein	154453	155169	-	ATG	TGA	716	238	PHIKZ147	*Pseudomonas* phage φKZ	2.8e-158
gp184	Virion associated protein	155159	155665	-	ATG	TAA	506	168	PHIKZ148	*Pseudomonas* phage φKZ	1.5e-106
gp185	β/β′-like virion-associated protein	155668	156342	-	ATG	TAA	674	224	PHIKZ149	*Pseudomonas* phage φKZ	3.2e-145
gp186	Uncharacterized protein	156375	156695	-	ATG	TAA	320	106	PHIKZ150	*Pseudomonas* phage φKZ	1.4e-66
gp187	Uncharacterized protein	156719	157069	-	ATG	TAA	350	116	PHIKZ151	*Pseudomonas* phage φKZ	2.0e-75
gp188	Virion associated protein	157146	158585	-	ATG	TAA	1439	479	PHIKZ152	*Pseudomonas* phage φKZ	0.0e+00
gp189	Structural head protein	158683	159597	+	ATG	TAA	914	304	PHIKZ153	*Pseudomonas* phage φKZ	2.4e-196
gp190	Uncharacterized protein	159643	160062	-	ATG	TGA	419	139	PHIKZ154	*Pseudomonas* phage φKZ	4.8e-92
gp191	Ribonuclease H	160073	161521	-	ATG	TAG	1448	482	PHIKZ155	*Pseudomonas* phage φKZ	0.0e+00
gp192	Uncharacterized protein	161577	161978	-	ATG	TAA	401	133	PHIKZ156	*Pseudomonas* phage φKZ	3.7e-84
gp193	Uncharacterized protein	161999	162289	-	ATG	TAA	290	96	PHIKZ156.1	*Pseudomonas* phage φKZ	5.1e-54
gp194	Structural head protein	162468	163805	+	ATG	TAA	1337	445	PHIKZ157	*Pseudomonas* phage φKZ	4.9e-304
gp195	Virion associated protein	163861	164811	+	ATG	TGA	950	316	PHIKZ158	*Pseudomonas* phage φKZ	6.8e-219
gp196	Virion associated protein	164808	165212	+	ATG	TAA	404	134	PHIKZ159	*Pseudomonas* phage φKZ	3.2e-87
gp197	Structural protein	165216	165689	+	ATG	TAA	473	157	PHIKZ160	*Pseudomonas* phage φKZ	1.8e-105
gp198	Virion associated protein	165788	166423	+	ATG	TAA	635	211	PHIKZ161	*Pseudomonas* phage φKZ	6.4e-128
gp199	Structural head protein	166459	168027	+	ATG	TAG	1568	522	PHIKZ162	*Pseudomonas* phage φKZ	0.0e+00
gp200	Structural head protein	168037	169224	+	ATG	TAA	1187	395	PHIKZ163	*Pseudomonas* phage φKZ	9.0e-253
gp201	Uncharacterized protein	169264	169530	-	ATG	TAA	266	88	PHIKZ163.1	*Pseudomonas* phage φKZ	1.6e-54
gp202	Uncharacterized protein	169540	170421	-	ATG	TAA	881	293	PHIKZ164	*Pseudomonas* phage φKZ	1.0e-195
gp203	SbcC protein	170486	172978	+	ATG	TAA	2492	830	PHIKZ165	*Pseudomonas* phage φKZ	0.0e+00
gp204	Uncharacterized protein	173021	173197	-	ATG	TAA	176	58	PHIKZ165.1	*Pseudomonas* phage φKZ	7.8e-31
gp205	Uncharacterized protein	173289	173684	+	ATG	TGA	395	131	PHIKZ166	*Pseudomonas* phage φKZ	8.3e-86
gp206	Uncharacterized protein	173826	174032	+	ATG	TAA	206	68	PHIKZ166.1	*Pseudomonas* phage φKZ	2.3e-31
gp207	Uncharacterized protein	174051	174479	+	ATG	TAA	428	142	PHIKZ167	*Pseudomonas* phage φKZ	4.5e-90
gp208	Uncharacterized protein	174529	175185	+	ATG	TAA	656	218	PHIKZ168	*Pseudomonas* phage φKZ	1.0e-136
gp209	Virion associated protein	175231	175746	-	ATG	TGA	515	171	PHIKZ169	*Pseudomonas* phage φKZ	1.1e-104
gp210	Uncharacterized protein	175826	176350	+	ATG	TGA	524	174	PHIKZ170	*Pseudomonas* phage φKZ	8.2e-110
gp211	Uncharacterized protein	176421	177182	+	ATG	TAA	761	253	PHIKZ171	*Pseudomonas* phage φKZ	2.3e-174

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp212	Uncharacterized protein	177175	177615	+	ATG	TAA	440	146	PHIKZ172	*Pseudomonas* phage φKZ	3.4e-95
gp213	Uncharacterized protein	177620	178546	+	ATG	TAG	926	308	PHIKZ173	*Pseudomonas* phage φKZ	3.0e-204
gp214	Structural protein	178630	179694	+	ATG	TGA	1064	354	PHIKZ174	*Pseudomonas* phage φKZ	4.0e-235
gp215	Head protease	179743	180555	-	ATG	TAA	812	270	PHIKZ175	*Pseudomonas* phage φKZ	2.8e-186
gp216	Structural head protein	180567	181289	-	ATG	TAA	722	270	PHIKZ176	*Pseudomonas* phage φKZ	1.6e-160
gp217	Structural head protein	181443	183002	-	ATG	TAA	1559	519	PHIKZ177	*Pseudomonas* phage φKZ	0.0e+00
gp218	Uncharacterized protein	183014	183211	-	ATG	TAA	197	65	PHIKZ177.1	*Pseudomonas* phage φKZ	3.1e-34
gp219	β/β′-like virion-associated protein	183195	187550	-	ATG	TAA	4355	1451	PHIKZ178	*Pseudomonas* phage φKZ	0.0e+00
gp220	β/β′-like virion-associated protein	187553	189205	-	ATG	TAA	1652	550	PHIKZ180	*Pseudomonas* phage φKZ	0.0e+00
gp221	Structural peptidoglycan hydrolase	189281	195994	+	ATG	TAA	6713	2237	PHIKZ181	*Pseudomonas* phage φKZ	0.0e+00
gp222	Structural protein	196050	198044	+	ATG	TAA	1994	664	PHIKZ182	*Pseudomonas* phage φKZ	0.0e+00
gp223	Uncharacterized protein	198103	198408	+	ATG	TGA	305	101	PHIKZ183	*Pseudomonas* phage φKZ	1.5e-56
gp224	Uncharacterized protein	198416	199108	+	ATG	TGA	692	230	PHIKZ184	*Pseudomonas* phage φKZ	4.5e-153
gp225	Uncharacterized protein	199217	199771	+	ATG	TAG	554	184	PHIKZ185	*Pseudomonas* phage φKZ	2.0e-116
gp226	Uncharacterized protein	199773	200015	+	ATG	TAA	242	80	PHIKZ185.1	*Pseudomonas* phage φKZ	9.8e-42
gp227	Uncharacterized protein	200046	200450	+	ATG	TAA	404	134	PHIKZ186	*Pseudomonas* phage φKZ	1.3e-79
gp228	Virion associated protein	200571	200900	+	ATG	TAG	329	109	PHIKZ187	*Pseudomonas* phage φKZ	2.0e-62
gp229	Thymidylate kinase	200903	201961	+	ATG	TAA	1058	352	PHIKZ188	*Pseudomonas* phage φKZ	1.0e-234
gp230	Uncharacterized protein	201999	202250	+	ATG	TAA	251	83	PHIKZ189	*Pseudomonas* phage φKZ	2.3e-49
gp231	Uncharacterized protein	202319	202759	+	ATG	TAG	440	146	PHIKZ190	*Pseudomonas* phage φKZ	1.1e-89
gp232	Virion associated protein	202787	203578	+	ATG	TAA	791	263	PHIKZ191	*Pseudomonas* phage φKZ	4.1e-175
gp233	Uncharacterized protein	203571	204296	+	ATG	TAA	725	241	PHIKZ192	*Pseudomonas* phage φKZ	5.7e-168
gp234	Uncharacterized protein	204296	204658	+	ATG	TAA	362	120	PHIKZ193	*Pseudomonas* phage φKZ	6.3e-77
gp235	Uncharacterized protein	204671	205027	+	ATG	TAA	356	118	PHIKZ194	*Pseudomonas* phage φKZ	1.4e-71
gp236	Uncharacterized protein	205177	205551	+	ATG	TAA	374	124	PHIKZ195	*Pseudomonas* phage φKZ	2.5e-78
gp237	Uncharacterized protein	205553	205915	+	ATG	TAG	362	120	PHIKZ196	*Pseudomonas* phage φKZ	1.9e-76
gp238	Uncharacterized protein	205927	206289	+	ATG	TAA	362	120	PHIKZ197	*Pseudomonas* phage φKZ	2.0e-76
gp239	Virion associated protein	206431	206919	+	ATG	TAA	488	162	PHIKZ198	*Pseudomonas* phage φKZ	3.3e-104
gp240	Structural protein	207010	207441	+	ATG	TAA	431	143	PHIKZ199	*Pseudomonas* phage φKZ	4.2e-93
gp241	Virion associated protein	207461	208120	+	ATG	TAA	659	219	PHIKZ200	*Pseudomonas* phage φKZ	1.0e-139

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp242	Structural protein	208161	210110	-	ATG	TAA	1949	649	PHIKZ201	*Pseudomonas* phage φKZ	0.0e+00
gp243	Structural protein	210121	210591	-	ATG	TAA	470	156	PHIKZ202	*Pseudomonas* phage φKZ	1.9e-96
gp244	Structural protein	210603	212714	-	ATG	TAA	2111	703	PHIKZ203	*Pseudomonas* phage φKZ	0.0e+00
gp245	Uncharacterized protein	213065	213460	+	ATG	TAA	395	131	PHIKZ204	*Pseudomonas* phage φKZ	1.3e-82
gp246	Uncharacterized protein	213432	213608	+	ATG	TAA	176	58	PHIKZ204.1	*Pseudomonas* phage φKZ	1.2e-31
gp247	Uncharacterized protein	213625	214026	+	ATG	TGA	401	133	PHIKZ205	*Pseudomonas* phage φKZ	3.2e-82
gp248	Virion associated protein	214269	214979	+	ATG	TAA	710	236	PHIKZ206	*Pseudomonas* phage φKZ	2.3e-162
gp249	Uncharacterized protein	214979	215632	+	ATG	TAA	653	217	PHIKZ207	*Pseudomonas* phage φKZ	2.5e-147
gp250	Uncharacterized protein	215736	216836	+	ATG	TAA	1100	366	PHIKZ208	*Pseudomonas* phage φKZ	1.3e-238
gp251	Uncharacterized protein	216944	217546	+	ATG	TGA	602	200	PHIKZ209	*Pseudomonas* phage φKZ	3.4e-132
gp252	Uncharacterized protein	217674	218006	+	ATG	TAG	332	110	PHIKZ210	*Pseudomonas* phage φKZ	2.1e-74
gp253	Uncharacterized protein	218018	218143	+	ATG	TAG	125	41	PHIKZ210.1	*Pseudomonas* phage φKZ	4.1e-22
gp254	Uncharacterized protein	218269	218526	+	ATG	TAA	257	85	PHIKZ210.2	*Pseudomonas* phage φKZ	1.3e-51
gp255	Uncharacterized protein	218539	218958	+	ATG	TGA	419	139	PHIKZ211	*Pseudomonas* phage φKZ	4.8e-88
gp256	Uncharacterized protein	218955	219299	+	ATG	TAA	344	114	PHIKZ212	*Pseudomonas* phage φKZ	9.9e-75
gp257	Uncharacterized protein	219313	219852	+	ATG	TGA	539	179	PHIKZ213	*Pseudomonas* phage φKZ	7.4e-125
gp258	Uncharacterized protein	219876	220445	+	ATG	TAA	569	189	PHIKZ214	*Pseudomonas* phage φKZ	3.8e-125
gp259	Uncharacterized protein	220448	220684	+	ATG	TGA	236	78	PHIKZ215	*Pseudomonas* phage φKZ	5.1e-47
gp260	Uncharacterized protein	220775	221251	+	ATG	TAA	476	158	PHIKZ216	*Pseudomonas* phage φKZ	5.2e-100
gp261	Uncharacterized protein	221309	221500	+	ATG	TAA	191	63	PHIKZ217	*Pseudomonas* phage φKZ	8.9e-28
gp262	Virion associated protein	221520	221735	+	ATG	TAA	215	71	PHIKZ218	*Pseudomonas* phage φKZ	4.1e-38
gp263	Uncharacterized protein	221846	222145	+	GTG	TAA	299	99	PHIKZ218.1	*Pseudomonas* phage φKZ	4.1e-64
gp264	Uncharacterized protein	222195	223328	+	ATG	TGA	1133	377	PHIKZ219	*Pseudomonas* phage φKZ	2.3e-249
gp265	Uncharacterized protein	223325	223795	+	ATG	TAA	470	156	PHIKZ220	*Pseudomonas* phage φKZ	2.5e-96
gp266	Uncharacterized protein	223898	224245	+	ATG	TGA	347	115	PHIKZ221	*Pseudomonas* phage φKZ	6.8e-69
gp267	Uncharacterized protein	224314	224646	+	ATG	TAA	332	110	PHIKZ222	*Pseudomonas* phage φKZ	5.6e-70
gp268	Uncharacterized protein	224835	225233	+	ATG	TAG	398	132	PHIKZ223	*Pseudomonas* phage φKZ	5.3e-88
gp269	Structural protein	225241	225654	+	ATG	TAA	413	137	PHIKZ224	*Pseudomonas* phage φKZ	1.2e-89
gp270	Uncharacterized protein	225777	225968	+	ATG	TGA	191	63	PHIKZ225	*Pseudomonas* phage φKZ	1.0e-26
gp271	Uncharacterized protein	225955	226284	+	ATG	TGA	329	109	PHIKZ226	*Pseudomonas* phage φKZ	4.8e-67
gp272	Uncharacterized protein	226294	226809	+	ATG	TAA	515	171	PHIKZ227	*Pseudomonas* phage φKZ	2.4e-113
gp273	Structural protein	226822	227421	+	ATG	TGA	599	199	PHIKZ228	*Pseudomonas* phage φKZ	5.5e-126
gp274	Uncharacterized protein	227478	227843	+	ATG	TAA	365	121	PHIKZ229	*Pseudomonas* phage φKZ	1.9e-75

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp275	Uncharacterized protein	227900	228196	+	ATG	TGA	296	98	PHIKZ230	*Pseudomonas* phage φKZ	8.5e-60
gp276	Uncharacterized protein	228189	228479	+	ATG	TAA	290	96	PHIKZ230.1	*Pseudomonas* phage φKZ	9.4e-57
gp277	Uncharacterized protein	228581	228754	+	ATG	TAA	173	57	PHIKZ230.2	*Pseudomonas* phage φKZ	6.7e-30
gp278	Uncharacterized protein	228754	229221	+	ATG	TAA	467	155	PHIKZ231	*Pseudomonas* phage φKZ	1.3e-99
gp279	Virion associated protein	229244	230152	+	ATG	TGA	908	302	PHIKZ232	*Pseudomonas* phage φKZ	1.5e-196
gp280	Uncharacterized protein	230149	230406	+	ATG	TAG	257	85	PHIKZ232.2	*Pseudomonas* phage φKZ	7.2e-27
gp281	Uncharacterized protein	230415	231122	+	ATG	TAA	707	235	PHIKZ233	*Pseudomonas* phage φKZ	2.2e-155
gp282	Uncharacterized protein	231228	231521	+	ATG	TGA	293	97	PHIKZ234	*Pseudomonas* phage φKZ	7.3e-60
gp283	Uncharacterized protein	231549	231749	+	ATG	TAA	200	66	PHIKZ234.1	*Pseudomonas* phage φKZ	7.1e-39
gp284	Thymidylate synthase	231749	233215	+	ATG	TAA	1466	488	PHIKZ235	*Pseudomonas* phage φKZ	0.0e+00
gp285	Uncharacterized protein	233318	233659	+	ATG	TAA	341	113	PHIKZ236	*Pseudomonas* phage φKZ	1.3e-43
gp286	Uncharacterized protein	233671	234183	+	ATG	TAA	512	170	Uncharacterized protein	*Pseudomonas* phage PhiPA3	2.4e-05
gp287	Uncharacterized protein	234339	234620	+	ATG	TGA	281	93			
gp288	Uncharacterized protein	234617	234787	+	ATG	TAG	170	56	PHIKZ237	*Pseudomonas* phage φKZ	2.2e-25
gp289	Uncharacterized protein	235025	235519	+	ATG	TAA	494	164	PHIKZ238	*Pseudomonas* phage φKZ	1.5e-46
gp290	Uncharacterized protein	235609	235902	+	ATG	TGA	293	97	PHIKZ238.1	*Pseudomonas* phage φKZ	2.0e-51
gp291	Uncharacterized protein	235974	236096	+	ATG	TAA	122	40	PHIKZ238.2	*Pseudomonas* phage φKZ	2.3e-18
gp292	Uncharacterized protein	236262	236747	+	ATG	TAA	485	161	PHIKZ240	*Pseudomonas* phage φKZ	4.3e-87
gp293	Uncharacterized protein	236877	237206	+	ATG	TAA	329	109	PHIKZ241	*Pseudomonas* phage φKZ	8.3e-51
gp294	Uncharacterized protein	237318	237755	+	ATG	TAA	437	145	PHIKZ242	*Pseudomonas* phage φKZ	1.7e-95
gp295	Uncharacterized protein	237776	238672	+	ATG	TAA	896	298	PHIKZ243	*Pseudomonas* phage φKZ	1.6e-183
gp296	Structural protein	238730	239086	+	ATG	TAA	356	118	PHIKZ244	*Pseudomonas* phage φKZ	3.5e-73
gp297	Uncharacterized protein	239137	239571	+	ATG	TAG	434	144	PHIKZ245	*Pseudomonas* phage φKZ	3.0e-95
gp298	Virion associated protein	239602	240258	+	ATG	TAA	656	218	PHIKZ246	*Pseudomonas* phage φKZ	2.3e-142
gp299	Uncharacterized protein	240378	240641	+	ATG	TAA	263	87	PHIKZ246.1	*Pseudomonas* phage φKZ	1.0e-54
gp300	Uncharacterized protein	240668	241075	+	ATG	TAA	407	135	PHIKZ247	*Pseudomonas* phage φKZ	1.1e-83
gp301	Uncharacterized protein	241088	241447	+	ATG	TAG	359	119	PHIKZ248	*Pseudomonas* phage φKZ	1.2e-73
gp302	Uncharacterized protein	241451	242245	+	ATG	TAA	794	264	PHIKZ249	*Pseudomonas* phage φKZ	3.2e-179
gp303	Uncharacterized protein	242259	242642	+	ATG	TAA	383	127	PHIKZ250	*Pseudomonas* phage φKZ	7.1e-76
gp304	Uncharacterized protein	242660	243115	+	ATG	TAA	455	151	PHIKZ251	*Pseudomonas* phage φKZ	6.5e-95
gp305	Uncharacterized protein	243148	243525	+	ATG	TAG	377	125	PHIKZ252	*Pseudomonas* phage φKZ	1.6e-75
gp306	Uncharacterized protein	243515	243754	+	ATG	TAA	239	79	PHIKZ253	*Pseudomonas* phage φKZ	7.1e-49

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp307	Uncharacterized protein	243837	243944	+	ATG	TAA	107	35	PHIKZ253.1	*Pseudomonas* phage φKZ	4.9e-17
gp308	Uncharacterized protein	244048	244422	+	ATG	TAA	374	124	PHIKZ254	*Pseudomonas* phage φKZ	6.3e-77
gp309	Uncharacterized protein	244409	244870	+	ATG	TAG	461	153	PHIKZ255	*Pseudomonas* phage φKZ	5.1e-98
gp310	Uncharacterized protein	244843	245073	+	ATG	TGA	230	76	PHIKZ256	*Pseudomonas* phage φKZ	8.6e-49
gp311	Uncharacterized protein	245070	245756	+	ATG	TAA	686	228	PHIKZ257	*Pseudomonas* phage φKZ	5.0e-152
gp312	Uncharacterized protein	245801	246175	+	ATG	TAG	374	124	PHIKZ258	*Pseudomonas* phage φKZ	1.7e-78
gp313	Uncharacterized protein	246178	246612	+	ATG	TAA	434	144	PHIKZ259	*Pseudomonas* phage φKZ	6.2e-81
gp314	Uncharacterized protein	246737	247096	+	ATG	TAA	359	119	PHIKZ260	*Pseudomonas* phage φKZ	8.1e-75
gp315	Uncharacterized protein	247142	247672	+	ATG	TAG	530	176	Uncharacterized protein	*Pseudomonas* phage PA7	1.0e-99
gp316	Uncharacterized protein	247693	248115	+	ATG	TAA	422	140	PHIKZ261	*Pseudomonas* phage φKZ	1.2e-84
gp317	Uncharacterized protein	248169	248402	+	ATG	TGA	233	77	PHIKZ262	*Pseudomonas* phage φKZ	4.8e-44
gp318	Uncharacterized protein	248430	248708	+	ATG	TAA	278	92	PHIKZ263	*Pseudomonas* phage φKZ	2.4e-52
gp319	Uncharacterized protein	248733	249152	+	ATG	TAA	419	139	PHIKZ264	*Pseudomonas* phage φKZ	1.2e-84
gp320	Uncharacterized protein	249231	249527	+	ATG	TGA	296	98	PHIKZ265	*Pseudomonas* phage φKZ	6.2e-65
gp321	Uncharacterized protein	249520	249957	+	GTG	TAA	437	145	Uncharacterized protein	*Pseudomonas* phage PA7	1.5e-88
gp322	Uncharacterized protein	249994	250425	+	ATG	TAA	431	143	PHIKZ266	*Pseudomonas* phage φKZ	3.3e-93
gp323	Uncharacterized protein	250488	250790	+	ATG	TAA	302	100	PHIKZ267	*Pseudomonas* phage φKZ	5.5e-63
gp324	Uncharacterized protein	250874	251026	+	ATG	TAG	152	50	PHIKZ267.1	*Pseudomonas* phage φKZ	2.8e-25
gp325	Uncharacterized protein	251083	251385	+	ATG	TGA	302	100	PHIKZ268	*Pseudomonas* phage φKZ	2.4e-63
gp326	Uncharacterized protein	251394	251606	+	ATG	TGA	212	70	PHIKZ269	*Pseudomonas* phage φKZ	5.9e-42
gp327	Uncharacterized protein	251603	251980	+	ATG	TGA	377	125	PHIKZ270	*Pseudomonas* phage φKZ	7.8e-78
gp328	Uncharacterized protein	251982	252347	+	ATG	TAA	365	121	PHIKZ271	*Pseudomonas* phage φKZ	3.6e-70
gp329	Uncharacterized protein	252391	252726	+	ATG	TAG	335	111	PHIKZ272	*Pseudomonas* phage φKZ	4.1e-68
gp330	Uncharacterized protein	252772	253458	+	ATG	TAA	686	228	PHIKZ273	*Pseudomonas* phage φKZ	2.4e-153
gp331	Uncharacterized protein	253471	253710	+	ATG	TGA	239	79	PHIKZ274	*Pseudomonas* phage φKZ	7.4e-49
gp332	Uncharacterized protein	253707	254066	+	ATG	TAA	359	119	PHIKZ275	*Pseudomonas* phage φKZ	2.5e-78
gp333	Uncharacterized protein	254081	254602	+	ATG	TAA	521	173	PHIKZ276	*Pseudomonas* phage φKZ	2.7e-111
gp334	Uncharacterized protein	254614	254913	+	ATG	TAA	299	99	PHIKZ277	*Pseudomonas* phage φKZ	6.3e-61
gp335	Uncharacterized protein	254972	255412	+	ATG	TAA	440	146	PHIKZ278	*Pseudomonas* phage φKZ	4.0e-95
gp336	Uncharacterized protein	255405	255917	+	ATG	TGA	512	170	PHIKZ279	*Pseudomonas* phage φKZ	3.1e-116
gp337	Uncharacterized protein	255931	256263	+	ATG	TAG	332	110	PHIKZ280	*Pseudomonas* phage φKZ	3.8e-70
gp338	Uncharacterized protein	256330	256686	+	ATG	TGA	356	118	PHIKZ281	*Pseudomonas* phage φKZ	1.7e-74
gp339	Virion associated protein	256733	257248	+	ATG	TAA	515	171	PHIKZ282	*Pseudomonas* phage φKZ	2.1e-99

Protein no.	Predicted function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Length (aa)	Homolog protein	Homolog bacteria	E-value
gp340	Uncharacterized protein	257264	257728	+	ATG	TAG	464	154	PHIKZ283	*Pseudomonas* phage φKZ	7.9e-98
gp341	Uncharacterized protein	257793	258146	+	GTG	TAA	353	117	PHIKZ283.1	*Pseudomonas* phage φKZ	4.7e-66
gp342	Virion associated protein	258188	259060	+	ATG	TAA	872	290	PHIKZ284	*Pseudomonas* phage φKZ	2.0e-172
gp343	Uncharacterized protein	259117	259902	+	ATG	TGA	785	261	PHIKZ285	*Pseudomonas* phage φKZ	1.0e-175
gp344	Uncharacterized protein	259899	261425	+	ATG	TAA	1526	508	PHIKZ286	*Pseudomonas* phage φKZ	0.0e+00
gp345	Uncharacterized protein	261449	261847	+	ATG	TAA	398	132	PHIKZ286.1	*Pseudomonas* phage φKZ	6.5e-22
gp346	Virion associated protein	261939	262592	+	ATG	TAA	653	217	PHIKZ287	*Pseudomonas* phage φKZ	8.7e-144
gp347	Uncharacterized protein	262594	263007	+	ATG	TAA	413	137	PHIKZ287.1	*Pseudomonas* phage φKZ	5.6e-28
gp348	Uncharacterized protein	263149	263529	+	ATG	TAA	380	126	PHIKZ288	*Pseudomonas* phage φKZ	7.9e-80
gp349	Uncharacterized protein	263696	264427	+	ATG	TGA	731	243	PHIKZ289	*Pseudomonas* phage φKZ	5.8e-161
gp350	Uncharacterized protein	264433	264948	+	ATG	TAA	515	171	PHIKZ290	*Pseudomonas* phage φKZ	1.6e-114
gp351	Uncharacterized protein	265117	265797	+	ATG	TAA	680	226	PHIKZ291	*Pseudomonas* phage φKZ	5.4e-139
gp352	Virion associated protein	265858	266247	-	ATG	TGA	389	129	PHIKZ292	*Pseudomonas* phage φKZ	1.0e-67
gp353	Uncharacterized protein	266328	266738	-	ATG	TAA	410	136	PHIKZ293	*Pseudomonas* phage φKZ	1.1e-78
gp354	Virion associated protein	266781	266957	-	ATG	TAA	176	58	PHIKZ293.1	*Pseudomonas* phage φKZ	5.2e-32
gp355	Uncharacterized protein	267307	267444	+	ATG	TAA	137	45	PHIKZ293.2	*Pseudomonas* phage φKZ	2.0e-19
gp356	Uncharacterized protein	267503	268129	-	ATG	TAA	626	208	PHIKZ294	*Pseudomonas* phage φKZ	5.5e-95
gp357	Uncharacterized protein	268294	268752	-	GTG	TAA	458	152	Virion structural protein	*Pseudomonas* phage 201phi2-1	5.0e-10
gp358	Virion associated protein	268974	269498	-	ATG	TAA	524	174	PHIKZ297	*Pseudomonas* phage φKZ	2.7e-66
gp359	Structural protein	269790	270797	-	ATG	TAA	1007	335	PHIKZ298	*Pseudomonas* phage φKZ	9.6e-225
gp360	Uncharacterized protein	271093	271440	-	ATG	TAA	347	115	PHIKZ299	*Pseudomonas* phage φKZ	4.9e-74
gp361	Uncharacterized protein	271443	271730	-	ATG	TAA	287	95	PHIKZ299.1	*Pseudomonas* phage φKZ	9.7e-60
gp362	Uncharacterized protein	271739	271960	-	ATG	TAG	221	73	PHIKZ300	*Pseudomonas* phage φKZ	1.1e-41
gp363	Uncharacterized protein	272097	272390	-	ATG	TAA	293	97	PHIKZ301	*Pseudomonas* phage φKZ	3.5e-63
gp364	Uncharacterized protein	272640	273092	-	ATG	TGA	452	150	PHIKZ302	*Pseudomonas* phage φKZ	1.4e-95
gp365	Structural protein	273318	275201	-	ATG	TAA	1883	627	PHIKZ303	*Pseudomonas* phage φKZ	1.7e-242
gp366	Uncharacterized protein	275287	275700	-	ATG	TAG	413	137	PHIKZ304	*Pseudomonas* phage φKZ	4.7e-81
gp367	Ribonucleoside reductase	275697	276860	-	ATG	TGA	1163	387	PHIKZ305	*Pseudomonas* phage φKZ	2.2e-258
gp368	Putative ribonucleotide-diphosphate reductase subunit alpha, NrdA	276961	279249	-	ATG	TAA	2288	762	PHIKZ306	*Pseudomonas* phage φKZ	0.0e+00

Table S 6.2: The genome annotations of giant phage PA5oct.

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp001	Uncharacterized protein	65	307	+	ATG	TAG	243	80		
gp002	Uncharacterized protein	424	1509	+	ATG	TAA	1086	361		
gp003	Uncharacterized protein	1551	1772	+	ATG	TAG	222	73		
gp004	Uncharacterized protein	1952	2275	+	ATG	TAA	324	107		
gp005	Uncharacterized protein	2278	2604	+	ATG	TAA	327	108		
gp006	Uncharacterized protein	2614	2871	+	ATG	TGA	258	85		
gp007	Uncharacterized protein	2868	3176	+	ATG	TGA	309	102		
gp008	Uncharacterized protein	3184	3669	+	ATG	TAA	486	161		
gp009	Uncharacterized protein	3673	3867	+	ATG	TGA	195	64		
tRNA		4169	4242	MET (CAT)			74	24		
tRNA		4268	4349	LEU (TAA)			82	27		
tRNA		4403	4475	ASN (GTT)			73	24		
tRNA		4650	4723	MET (CAT)			74	24		
tRNA		4777	4858	LEU (TAA)			82	27		
gp010	Uncharacterized protein	4890	5078	+	ATG	TGA	189	62		
gp011	Uncharacterized protein	5057	5251	+	ATG	TAA	195	64		
gp012	Uncharacterized protein	5374	5523	+	ATG	TGA	150	49		
gp013	Uncharacterized protein	5520	5714	+	ATG	TAG	195	64		
tRNA		5817	5890	ARG (TCT)			74	24		
gp014	Uncharacterized protein	5860	6144	+	ATG	TAA	285	94		
gp015	Uncharacterized protein	6162	6404	+	ATG	TAA	243	80		
gp016	Uncharacterized protein	6433	6765	+	ATG	TAA	333	110	Putative uncharacterized protein gp70 *Pseudomonas* phage PAJU2	8,00E-47
tRNA		6875	6948	PRO (TGG)			74	24		
tRNA		6961	7033	GLY (TCC)			73	24		
gp017	Uncharacterized protein	7070	7237	+	ATG	TAG	168	55		
gp018	Uncharacterized protein	7279	7467	+	ATG	TAG	189	62		
gp019	Uncharacterized protein	7582	7749	+	ATG	TGA	168	55		
gp020	Structural protein	7839	8198	+	ATG	TAG	360	119		
tRNA		8213	8299	SER (TGA)			87	28		
tRNA		8308	8394	SER (GCT)			87	28		
gp021	Uncharacterized protein	8724	8939	+	ATG	TAA	216	71		
gp022	Uncharacterized protein	8944	9162	+	ATG	TAA	219	72		
gp023	Uncharacterized protein	9153	9425	+	ATG	TGA	273	90		
gp024	Uncharacterized protein	9584	9934	+	ATG	TAA	351	116		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp025	Uncharacterized protein	9935	10114	+	ATG	TGA	180	59		
gp026	Uncharacterized protein	10111	10329	+	ATG	TGA	219	72		
gp027	Uncharacterized protein	10326	10844	+	ATG	TAA	519	173	*Alteromonas* phage vB_AmaP_AD45-P2 hypothetical protein	8,00E-38
gp028	Uncharacterized protein	10906	11544	+	ATG	TAA	639	212		
gp029	Uncharacterized protein	11651	11881	+	ATG	TGA	231	76		
gp030	Uncharacterized protein	11878	12123	+	ATG	TAG	246	81		
gp031	Uncharacterized protein	12245	12433	+	ATG	TAA	189	62		
gp032	Uncharacterized protein	12519	12632	+	ATG	TAG	114	37		
gp033	Uncharacterized protein	12964	13185	+	ATG	TAA	222	73		
gp034	Uncharacterized protein	13196	13372	+	ATG	TAA	177	58		
gp035	Uncharacterized protein	13360	13656	+	ATG	TAA	297	98		
gp036	Uncharacterized protein	13660	13959	+	ATG	TAG	300	99		
gp037	Uncharacterized protein	13962	14573	+	TTG	TAA	612	203		
gp038	Uncharacterized protein	14566	14685	+	ATG	TGA	120	39		
gp039	Uncharacterized protein	14682	15032	+	ATG	TAA	351	116		
gp040	Uncharacterized protein	15244	15576	+	ATG	TAA	333	110		
gp041	Uncharacterized protein	15625	15966	+	ATG	TAG	342	113		
gp042	Uncharacterized protein	15951	16298	+	ATG	TAG	348	115		
gp043	Uncharacterized protein	16288	16473	+	ATG	TAG	186	61		
gp044	Uncharacterized protein	16448	16771	+	ATG	TGA	324	99		
gp045	Murein hydrolase,	16768	17262	+	ATG	TAA	495	164		
gp046	Peptidase M23	17255	17596	+	ATG	TAA	342	113		
gp047	Uncharacterized protein	17634	17921	+	ATG	TGA	288	95		
gp048	Uncharacterized protein	18034	18327	+	ATG	TAA	294	97		
gp049	Uncharacterized protein	18356	18649	+	ATG	TGA	294	97		
gp050	Uncharacterized protein	18732	18878	+	ATG	TAA	147	48		
gp051	Uncharacterized protein	18871	19131	+	ATG	TGA	261	86		
gp052	Uncharacterized protein	19163	19456	+	ATG	TGA	294	97		
gp053	Uncharacterized protein	19453	19734	+	ATG	TAA	282	93		
gp054	Uncharacterized protein	19804	20109	+	ATG	TAA	306	101		
gp055	Uncharacterized protein	20097	20378	+	ATG	TAA	282	93		
gp056	Uncharacterized protein	20394	20702	+	ATG	TAG	309	102		
gp057	Uncharacterized protein	20723	20974	+	ATG	TAA	252	83		
gp058	Uncharacterized protein	20987	21376	+	ATG	TGA	390	129		
gp059	Uncharacterized protein	21360	21644	+	ATG	TGA	285	94		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp060	Uncharacterized protein	21641	21895	+	ATG	TAA	255	84		
gp061	Uncharacterized protein	21898	22179	+	ATG	TAA	282	93		
gp062	Uncharacterized protein	22166	22450	+	ATG	TGA	285	94		
gp063	Uncharacterized protein	22447	22731	+	ATG	TGA	285	94		
gp064	Uncharacterized protein	22728	23009	+	ATG	TAA	282	93		
gp065	Uncharacterized protein	23550	23819	+	ATG	TAA	270	89		
gp066	Uncharacterized protein	23823	24131	+	ATG	TAG	309	102		
gp067	Uncharacterized protein	24145	24429	+	ATG	TAG	285	94		
gp068	Uncharacterized protein	24552	24707	+	ATG	TAG	156	51		
gp069	Uncharacterized protein	24735	24983	+	ATG	TGA	249	82		
gp070	Uncharacterized protein	24989	25279	+	ATG	TAA	291	96		
gp071	Uncharacterized protein	25272	25394	+	ATG	TAA	123	40		
gp072	Uncharacterized protein	25397	25510	+	ATG	TAA	114	37		
gp073	Uncharacterized protein	25524	25814	+	ATG	TAA	291	96		
gp074	Uncharacterized protein	25814	26110	+	ATG	TAA	297	98		
gp075	Uncharacterized protein	26110	26265	+	ATG	TGA	156	51		
gp076	Uncharacterized protein	26250	26369	+	ATG	TAG	120	39		
gp077	Uncharacterized protein	26518	26649	+	ATG	TAA	132	43		
gp078	Uncharacterized protein	26728	26853	+	ATG	TAA	126	41		
gp079	Uncharacterized protein	26903	27037	+	ATG	TAA	135	44		
gp080	Uncharacterized protein	27012	27176	+	ATG	TAA	165	54		
gp081	Uncharacterized protein	27157	27306	+	ATG	TAA	150	49		
gp082	Uncharacterized protein	27312	27446	+	ATG	TAA	135	44		
gp083	Uncharacterized protein	27433	27588	+	ATG	TGA	156	51		
gp084	Uncharacterized protein	27585	27743	+	ATG	TGA	159	52		
gp085	Uncharacterized protein	27740	27877	+	ATG	TAA	138	45		
gp086	Uncharacterized protein	27915	28049	+	ATG	TAA	135	44		
gp087	Uncharacterized protein	28158	28304	+	ATG	TAA	147	48		
gp088	Uncharacterized protein	28304	28444	+	ATG	TAG	141	46		
gp089	Uncharacterized protein	28438	28611	+	ATG	TAA	174	57		
gp090	Uncharacterized protein	28729	28881	+	ATG	TAA	153	50		
gp091	Uncharacterized protein	29029	29178	+	ATG	TAA	150	49		
gp092	Uncharacterized protein	29217	29360	+	ATG	TGA	144	47		
gp093	Uncharacterized protein	29362	29532	+	ATG	TAA	171	56		
gp094	Uncharacterized protein	29507	29698	+	ATG	TGA	192	63		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp095	Uncharacterized protein	29695	29883	+	ATG	TGA	189	62		
gp096	Uncharacterized protein	29880	30191	+	ATG	TAA	312	103		
gp097	Uncharacterized protein	30313	30420	+	ATG	TAG	108	35		
gp098	Uncharacterized protein	30672	31079	+	ATG	TAA	408	135		
gp099	Uncharacterized protein	31616	31795	+	ATG	TAA	180	59		
gp100	Structural protein	32238	32942	+	ATG	TAA	705	234		
gp101	Uncharacterized protein	33070	33237	+	ATG	TGA	168	55		
gp102	Uncharacterized protein	33197	33370	+	ATG	TAA	174	57		
gp103	Uncharacterized protein	33372	33521	+	ATG	TAA	150	49		
gp104	Uncharacterized protein	33708	33911	+	ATG	TAA	204	67		
gp105	Uncharacterized protein	33892	34158	+	ATG	TAG	267	88		
gp106	Uncharacterized protein	34162	34296	+	ATG	TGA	135	44		
gp107	Uncharacterized protein	34293	34424	+	ATG	TAA	132	43		
gp108	Uncharacterized protein	34417	34629	+	ATG	TAA	213	70		
gp109	Uncharacterized protein	34645	34986	+	ATG	TAA	342	113		
gp110	Uncharacterized protein	34988	35197	+	ATG	TAA	210	69		
gp111	Uncharacterized protein	35218	35394	+	ATG	TAA	177	58		
gp112	Structural protein	35408	35605	+	ATG	TAA	198	65		
gp113	Uncharacterized protein	35607	35864	+	ATG	TAA	258	85		
gp114	Uncharacterized protein	36781	37368	+	ATG	TAA	588	195		
gp115	Structural protein	37372	37920	+	ATG	TAA	549	182		
gp116	Uncharacterized protein	38079	38636	+	ATG	TAA	558	185	*Pseudomonas* phage NP3 hypothetical protein	5,00E-08
gp117	Uncharacterized protein	38726	38977	+	ATG	TGA	252	83	*Pseudomonas* phage PAJU2 hypothetical protein PAJU2_gp39	2,00E-05
gp118	Uncharacterized protein	38974	39240	+	ATG	TGA	267	88	*Pseudomonas* phage YMC11/02/R656 hypothetical protein	4,00E-29
gp119	Uncharacterized protein	39237	39533	+	ATG	TAG	297	98		
gp120	Uncharacterized protein	39681	40082	+	ATG	TGA	402	133	*Ralstonia* phage RSP15 hypothetical protein	7,00E-10
gp121	Uncharacterized protein	40086	40352	+	ATG	TGA	267	88		
gp122	Uncharacterized protein	40443	40544	+	ATG	TGA	102	33		
gp123	Uncharacterized protein	40541	40855	+	ATG	TAA	315	104		
gp124	Uncharacterized protein	40880	41107	+	ATG	TAA	228	75	*Pseudomonas* phage PAK_P5 hypothetical protein PAK_P500131	2,00E-42
gp125	Uncharacterized protein	41191	41421	+	ATG	TAA	231	76		
gp126	Uncharacterized protein	41634	41810	+	ATG	TAA	177	58		
gp127	Uncharacterized protein	42052	42393	+	ATG	TAG	342	113		
gp128	Uncharacterized protein	42421	42741	+	ATG	TAA	321	106		
gp129	Uncharacterized protein	42871	43095	+	ATG	TAA	225	74	*Pseudomonas* phage VCM hypothetical protein VCM_00001	1,00E-11

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp130	Uncharacterized protein	43103	43408	+	ATG	TAA	306	101		
gp131	Uncharacterized protein	43410	43703	+	ATG	TGA	294	97		
gp132	Uncharacterized protein	43939	44229	+	ATG	TGA	291	96		
gp133	Uncharacterized protein	44226	44414	+	ATG	TGA	189	62		
gp134	Uncharacterized protein	44415	44612	+	ATG	TAG	198	65		
gp135	Uncharacterized protein	44999	45142	+	ATG	TAG	144	47		
gp136	Uncharacterized protein	45272	45439	+	ATG	TGA	168	55		
gp137	Uncharacterized protein	45436	45711	+	ATG	TGA	276	91		
gp138	Uncharacterized protein	45708	45926	+	ATG	TAA	219	72		
gp139	Uncharacterized protein	45937	46143	+	ATG	TAA	207	68		
gp140	Uncharacterized protein	46159	46362	+	ATG	TAA	204	67		
gp141	Uncharacterized protein	46373	46567	+	ATG	TAG	195	64		
gp142	Uncharacterized protein	46593	46892	+	ATG	TAA	300	99		
gp143	Uncharacterized protein	46914	47018	+	ATG	TAG	105	34		
gp144	Uncharacterized protein	47035	47178	+	ATG	TGA	144	47		
gp145	Uncharacterized protein	47175	47390	+	ATG	TAA	216	71		
gp146	Uncharacterized protein	47464	47661	+	ATG	TGA	198	65		
gp147	Uncharacterized protein	47840	48292	+	ATG	TAA	453	150		
gp148	Uncharacterized protein	48385	49032	+	ATG	TAA	648	215		
gp149	Uncharacterized protein	49034	49261	+	ATG	TAG	228	75		
gp150	Uncharacterized protein	49304	49453	+	ATG	TGA	150	49		
gp151	Uncharacterized protein	49610	49720	+	ATG	TAA	111	36		
gp152	Uncharacterized protein	49696	50040	+	ATG	TAA	345	114		
gp153	Uncharacterized protein	50597	51064	+	ATG	TAA	468	155		
gp154	Uncharacterized protein	51219	51356	+	ATG	TAG	138	45		
gp155	Uncharacterized protein	51371	51661	+	ATG	TAA	291	96		
gp156	Uncharacterized protein	51817	52065	+	ATG	TAA	249	82		
gp157	Uncharacterized protein	52077	52379	+	ATG	TAA	303	100		
gp158	Uncharacterized protein	52593	52913	+	ATG	TAG	321	106		
gp159	Uncharacterized protein	53310	53615	+	ATG	TAA	306	101		
gp160	Uncharacterized protein	53752	54039	+	ATG	TAA	288	95		
gp161	Uncharacterized protein	54221	54592	+	ATG	TAA	372	123		
gp162	Uncharacterized protein	54599	54886	+	ATG	TGA	288	95		
gp163	Uncharacterized protein	55037	55219	+	ATG	TAA	183	60		
gp164	Uncharacterized protein	55263	55502	-	ATG	TAA	240	79		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp165	Uncharacterized protein	55696	55866	-	ATG	TAA	171	56		
gp166	Uncharacterized protein	55856	56236	+	ATG	TAA	381	126		
gp167	Uncharacterized protein	56265	56432	+	ATG	TGA	168	55		
gp168	Uncharacterized protein	56511	56882	+	ATG	TAA	372	123		
gp169	Uncharacterized protein	56863	57189	+	TTG	TGA	327	108		
gp170	Uncharacterized protein	57182	57307	+	ATG	TAA	126	41		
gp171	Uncharacterized protein	57522	57983	+	ATG	TAA	462	153	*Vibrio* phage CHOED hypothetical protein	2,00E-20
gp172	Uncharacterized protein	58212	58628	+	ATG	TGA	417	138		
gp173	Uncharacterized protein	58701	58934	+	ATG	TAG	234	77		
gp174	Uncharacterized protein	58985	59086	+	ATG	TGA	102	33		
gp175	Uncharacterized protein	59201	59524	+	ATG	TAG	324	107		
gp176	Uncharacterized protein	59553	59744	+	ATG	TAA	192	63		
gp177	Uncharacterized protein	59831	59947	+	ATG	TAA	117	38	*Pseudomonas* phage KPP23 hypothetical protein KPP23_055	1,00E-04
gp178	Uncharacterized protein	60019	60324	+	ATG	TAA	306	101		
gp179	Uncharacterized protein	60266	60472	+	ATG	TAG	207	68		
gp180	Uncharacterized protein	60426	60557	+	ATG	TAA	132	43		
gp181	Uncharacterized protein	60607	60816	+	ATG	TAA	210	69		
gp182	Uncharacterized protein	60850	61341	+	ATG	TAA	492	163	*Pseudomonas* phage PMG1 unnamed protein product	7,00E-85
gp183	Uncharacterized protein	61419	61667	+	ATG	TAA	249	82		
gp184	Uncharacterized protein	61808	62008	+	ATG	TAA	201	66		
gp185	Uncharacterized protein	62039	62203	+	ATG	TAA	165	54	*Pseudomonas* phage PAK_P2 hypothetical protein PAK_P200121	3,00E-17
gp186	Uncharacterized protein	62324	62620	+	ATG	TAA	297	98		
gp187	Uncharacterized protein	63281	63445	+	ATG	TAA	165	54		
gp188	Uncharacterized protein	63452	63661	+	ATG	TAG	210	69		
gp189	Uncharacterized protein	63639	63806	+	ATG	TAA	168	55		
gp190	Uncharacterized protein	63799	64014	+	ATG	TGA	216	71		
gp191	Uncharacterized protein	64001	64192	+	ATG	TGA	192	63		
gp192	Uncharacterized protein	64189	64380	+	ATG	TAG	192	63		
gp193	Uncharacterized protein	64373	64600	+	ATG	TAA	228	75		
gp194	Uncharacterized protein	64924	65298	+	ATG	TGA	375	124	*Escherichia* phage vB_EcoP_PhAPEC7 hypothetical protein	4,00E-45
gp195	Uncharacterized protein	65295	65510	+	ATG	TAG	216	71		
gp196	Uncharacterized protein	65510	65725	+	ATG	TAG	216	71		
gp197	Uncharacterized protein	65716	65910	+	ATG	TAA	195	64		
gp198	Uncharacterized protein	65901	66158	+	ATG	TAA	258	85		
gp199	Uncharacterized protein	66151	66372	+	ATG	TGA	222	73		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp200	Uncharacterized protein	66369	66608	+	ATG	TAA	240	79		
gp201	Uncharacterized protein	66748	66933	+	ATG	TAA	186	61		
gp202	Uncharacterized protein	67035	67220	+	ATG	TAA	186	61		
gp203	Uncharacterized protein	67224	68201	+	ATG	TAA	978	325		
gp204	Structural protein	68215	69435	-	ATG	TAA	1221	406		
gp205	Structural protein	69435	70871	-	ATG	TAA	1437	478	*Campylobacter* phage CP21 hypothetical protein	3,00E-12
gp206	Structural protein	70874	72619	-	ATG	TAA	1746	581	*Salmonella* phage 7-11 putative structural protein	2,00E-13
gp207	Structural protein	72623	73228	-	ATG	TAA	606	201	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	3,00E-25
gp208	Uncharacterized protein	73230	73355	-	ATG	TAA	126	41		
gp209	Structural protein	73380	84956	-	ATG	TAG	11577	3858	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	9,00E-85
gp210	Baseplate wedge	84963	88922	-	ATG	TAA	3960	1319	*Klebsiella* phage K64-1 hypothetical protein	1,00E-
gp211	Baseplate protein	88937	89329	-	ATG	TAA	393	130	*Cronobacter* phage vB_CsaM_GAP32 base plate protein	2,00E-12
gp212	Uncharacterized protein	89466	89708	+	ATG	TAA	243	80		
gp213	Uncharacterized protein	89753	90037	+	ATG	TAA	285	94		
gp214	Baseplate hub subunit and	90045	91478	-	ATG	TAA	1434	477	*Cronobacter* phage vB_CsaM_GAP32 baseplate hub subunit and	1,00E-48
gp215	Structural protein	91478	93502	-	ATG	TAA	2025	674	*Klebsiella* phage K64-1 hypothetical protein	2,00E-08
gp216	Uncharacterized protein	93502	93801	-	ATG	TGA	300	99	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	3,00E-10
gp217	Structural protein	93801	94436	-	ATG	TAA	636	211	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	6,00E-26
gp218	RNA polymerase Sigma	94539	95381	+	ATG	TAA	843	280	*Escherichia* phage PBECO 4 RNA polymerase sigma factor	5,00E-47
gp219	Endonuclease subunit	95393	96406	+	ATG	TAA	1014	337	*Enterobacteria* phage vB_KleM-RaK2 endonuclease subunit	2,00E-
gp220	Recombination	96406	98160	+	ATG	TAA	1755	584	*Cronobacter* phage vB_CsaM_GAP32 recombination	2,00E-
gp221	Structural protein	98173	98922	-	ATG	TAA	750	249	*Pseudomonas* phage EL hypothetical protein PPEV_gp075	2,00E-12
gp222	Uncharacterized protein	98925	99422	-	ATG	TAA	498	165		
gp223	Uncharacterized protein	99855	100007	+	ATG	TAA	153	50		
gp224	Uncharacterized protein	100153	100272	+	ATG	TAA	120	39		
gp225	Uncharacterized protein	100936	101388	+	ATG	TAA	453	150		
gp226	Uncharacterized protein	101612	102466	-	ATG	TAG	855	284	*Escherichia* phage 121Q baseplate hub subunit	8,00E-32
gp227	Structural protein	102536	104932	+	ATG	TAA	2397	798		
gp228	Portal vertex protein of	104962	106677	+	ATG	TAA	1716	571	*Cronobacter* phage vB_CsaM_GAP32 portal vertex protein of	5,00E-
gp229	Structural protein	106677	106898	+	ATG	TAA	222	73		
gp230	Structural protein	107004	107816	+	ATG	TAA	813	270		
gp231	Prohead core scaffolding	107835	108467	+	ATG	TAG	633	210	*Escherichia* phage 121Q capsid maturation protease	2,00E-59
gp232	Uncharacterized protein	108480	109586	+	ATG	TAA	1107	368	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	7,00E-45
gp233	Precursor of major head	109606	110793	+	ATG	TAA	1188	395	*Cronobacter* phage vB_CsaM_GAP32 precursor of major head	0.0
gp234	Uncharacterized protein	110912	112489	+	ATG	TAA	1578	525	*Klebsiella* phage K64-1 hypothetical protein	1,00E-53

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp235	Uncharacterized protein	112541	113005	+	ATG	TAA	465	154		
gp236	Uncharacterized protein	113007	113705	+	ATG	TAA	699	232		
gp237	Uncharacterized protein	113714	114019	+	ATG	TAA	306	101	*Pseudomonas* phage VCM hypothetical protein VCM_00075	9,00E-30
gp238	Phage tail fiber protein H	114080	117106	+	ATG	TAA	3027	1008	*Escherichia* ABU 83972 putative phage tail fiber protein H	2,00E-35
gp239	MutT/NUDIX hydrolase	117714	118031	-	ATG	TAA	318	105	*Escherichia* phage 121Q MutT-like protein	2,00E-28
gp240	Tail sheath stabilizer and	118141	119511	-	ATG	TAA	1371	456	*Cronobacter* phage vB_CsaM_GAP32 tail sheath stabilizer and	6,00E-41
gp241	Structural protein	119511	120602	-	ATG	TAA	1092	363	*Escherichia* phage PBECO 4 hypothetical protein	7,00E-31
gp242	Uncharacterized protein	120669	121166	-	ATG	TAA	498	165		
gp243	Tail fiber protein	121180	122715	-	ATG	TAA	1536	511	*Vibrio* phage CHOED tail fiber repeat family protein	2,00E-04
gp244	Structural protein	122717	124315	-	ATG	TAA	1599	532		
gp245	Structural protein	124315	126678	-	ATG	TAA	2364	787		
gp246	Uncharacterized protein	127262	127429	+	ATG	TAA	168	55		
gp247	Structural protein	127467	127877	+	ATG	TAA	411	136		
gp248	Structural protein	127807	128124	+	ATG	TAA	318	105		
gp249	Uncharacterized protein	128121	128384	-	ATG	TAA	264	87		
gp250	Glycosyl hydrolase,	128363	128878	-	ATG	TAA	516	171		
gp251	Uncharacterized protein	129000	129431	+	ATG	TAG	432	143		
gp252	Structural protein	129433	129645	+	ATG	TAA	213	70		
gp253	Uncharacterized protein	129645	130115	+	ATG	TGA	471	156		
gp254	Uncharacterized protein	130112	130357	+	ATG	TAA	246	81	*Vibrio* phage nt-1 putative permease	7,00E-08
gp255	Uncharacterized protein	130350	130619	+	ATG	TAA	270	89		
gp256	Uncharacterized protein	130620	131174	+	ATG	TAA	555	184		
gp257	Structural protein	131190	131753	+	ATG	TAA	564	187		
gp258	Uncharacterized protein	131825	132076	+	ATG	TAA	252	83		
gp259	Uncharacterized protein	132227	132757	+	ATG	TGA	531	176		
gp260	Uncharacterized protein	132741	132920	+	ATG	TGA	180	59		
gp261	Uncharacterized protein	132917	133204	+	ATG	TAA	288	95		
gp262	Uncharacterized protein	133255	133419	+	ATG	TAA	165	54		
gp263	Uncharacterized protein	133747	135930	+	ATG	TAA	2184	727	*Cronobacter* phage S13 putative rIIA protector from prophage-	6,00E-54
gp264	RIIB protector from	135923	136666	+	ATG	TAA	744	247	*Enterobacter* phage PG7 protector from prophage-induced early	2,00E-56
gp265	Uncharacterized protein	136750	138771	+	ATG	TAA	2022	673		
gp266	Uncharacterized protein	138919	139215	+	ATG	TAG	297	98		
gp267	Uncharacterized protein	139352	139543	+	ATG	TGA	192	63		
gp268	Uncharacterized protein	139540	139692	+	ATG	TGA	153	50		
gp269	Uncharacterized protein	139741	140859	+	ATG	TAA	1119	372	*Stenotrophomonas* phage vB_SmaS-DLP_6 hypothetical protein	3,00E-23

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp270	Neck protein	140895	141647	+	ATG	TAA	753	250	*Cronobacter* phage vB_CsaM_GAP32 neck protein	1,00E-49
gp271	Structural protein	141652	142293	+	ATG	TAA	642	213	*Enterobacteria* phage vB_KleM-RaK2 putative deoxynucleoside	5,00E-42
gp272	Tail sheath monomer	142411	145050	+	ATG	TAA	2640	879	*Escherichia* phage 121Q tail sheath protein	2,00E-
gp273	Homing endonuclease	145152	146768	+	ATG	TAA	1617	538	*Yersinia* phage phiR1-37 g171	7,00E-52
gp274	Structural protein	146857	147477	+	ATG	TAA	621	206	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	7,00E-38
gp275	Structural protein	147537	149015	+	ATG	TAA	1479	492	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	3,00E-10
gp276	Structural protein	149015	149470	+	ATG	TAA	456	151		
gp277	Head completion protein	149470	149898	+	ATG	TAA	429	142	*Enterobacteria* phage vB_KleM-RaK2 head completion protein	1,00E-32
gp278	Uncharacterized protein	149900	150406	+	ATG	TAA	507	168		
gp279	Structural protein	150431	151033	+	ATG	TAA	603	200		
gp280	Structural protein	151063	152568	+	ATG	TAA	1506	501		
gp281	DNA terminase packaging	152597	154192	+	ATG	TAA	1596	531	*Cronobacter* phage vB_CsaM_GAP32 DNA terminase packaging	1,00E-
gp282	Structural protein	154224	155714	+	ATG	TAA	1491	496		
gp283	Structural protein	155743	157233	+	TTG	TAA	1491	496	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	4,00E-28
gp284	Structural protein	157432	158418	+	ATG	TAA	987	328	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	1,00E-69
gp285	RecA-like protein	158525	158770	+	ATG	TAA	246	81	*Cronobacter* phage vB_CsaM_GAP32 RecA-like protein	2,00E-15
gp286	Uncharacterized protein	158823	160256	+	ATG	TAA	1434	477	*Vibrio* phage vB_VmeM-32 hypothetical protein VmeM32_00088	3,00E-04
gp287	RecA-like protein	160260	161531	+	ATG	TAA	1272	423	*Cronobacter* phage vB_CsaM_GAP32 RecA-like protein	6,00E-
gp288	Uncharacterized protein	161524	163359	+	ATG	TGA	1836	611	*Campylobacter* phage NCTC12673 Hef59	2,00E-46
gp289	Uncharacterized protein	163356	163556	+	ATG	TAA	201	66		
gp290	Uncharacterized protein	163546	163872	+	ATG	TAA	327	108		
gp291	Uncharacterized protein	163856	164320	+	ATG	TAA	465	154	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	2,00E-36
gp292	RNA-DNA + DNA-DNA	164333	165784	+	ATG	TAA	1452	483	*Cronobacter* phage vB_CsaM_GAP32 RNA-DNA + DNA-DNA	2,00E-
gp293	Uncharacterized protein	165786	166037	+	ATG	TAA	252	83		
gp294	Uncharacterized protein	166048	166551	+	TTG	TAA	504	167	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	2,00E-24
gp295	Uncharacterized protein	166551	167018	+	ATG	TAA	468	155		
gp296	Uncharacterized protein	167068	167739	+	ATG	TAA	672	223	*Escherichia* phage 121Q hypothetical protein PBI_121Q_14	7,00E-30
gp297	Uncharacterized protein	167741	168343	+	ATG	TGA	603	200	*Escherichia* phage PBECO 4 hypothetical protein	1,00E-37
gp298	Structural protein	168788	169657	+	ATG	TAA	870	289	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	3,00E-06
gp299	DNA primase	169693	170715	+	ATG	TAG	1023	340	*Cronobacter* phage vB_CsaM_GAP32 putative DNA primase	3,00E-43
gp300	DNA primase-helicase	170717	172156	+	ATG	TAA	1440	479	*Klebsiella* phage K64-1 DNA primase-helicase subunit	2,00E-76
gp301	Uncharacterized protein	172197	172436	+	ATG	TAA	240	79	*Enterobacteria* phage vB_KleM-RaK2 hypothetical protein	4,00E-06
gp302	Structural protein	172439	173050	+	ATG	TAA	612	203		
gp303	Structural protein	173043	174242	+	ATG	TAA	1200	399		
gp304	Uncharacterized protein	174327	175667	+	ATG	TAG	1341	446		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp305	Uncharacterized protein	175668	176687	+	ATG	TAG	1020	339		
gp306	Structural protein	176678	176923	+	ATG	TAA	246	81		
gp307	Structural protein	176923	177636	+	ATG	TAG	714	237		
gp308	Structural protein	177644	178075	+	ATG	TAG	432	143		
gp309	Structural protein	178142	178444	+	ATG	TAA	303	100		
gp310	Structural protein	178563	179027	+	ATG	TGA	465	154		
gp311	Structural protein	179024	179650	+	ATG	TAA	627	208		
gp312	Uncharacterized protein	179652	180107	+	ATG	TAA	456	151		
gp313	Structural protein	180144	181331	+	ATG	TAA	1188	395		
gp314	Structural protein	181332	181646	+	ATG	TAA	315	104		
gp315	Structural protein	181646	181981	+	ATG	TAA	336	111		
gp316	Exonuclease	182045	182710	+	ATG	TAA	666	221	*Escherichia* phage 121Q exonuclease	5,00E-50
gp317	Structural protein	182742	183941	+	ATG	TAA	1200	399		
gp318	Uncharacterized protein	183958	184392	+	ATG	TAA	435	144		
gp319	Structural protein	184401	186686	+	ATG	TAA	2286	761		
gp320	Structural protein	186740	187879	+	ATG	TAA	1140	379	*Caulobacter* phage Cr30 baseplate hub subunit and tail lysozyme	6,00E-27
gp321	Baseplate hub sub and tail	187904	188926	+	ATG	TAA	1023	340		
gp322	Uncharacterized protein	188969	189193	+	ATG	TAA	225	74		
gp323	Uncharacterized protein	189186	189401	+	ATG	TGA	216	71		
gp324	Uncharacterized protein	189495	190220	+	ATG	TGA	726	241		
gp325	Uncharacterized protein	190217	190585	+	ATG	TGA	369	122		
gp326	Uncharacterized protein	190578	190898	+	ATG	TAA	321	106		
gp327	Structural protein	190898	191794	+	ATG	TAA	897	298		
gp328	Uncharacterized protein	191794	192105	+	ATG	TAA	312	103		
gp329	GTP cyclohydrolase	192254	192943	+	ATG	TAA	690	229	*Vibrio* phage KVP40 GTP cyclohydrolase I	3,00E-83
gp330	Structural protein	193011	193529	+	TTG	TAA	519	172		
gp331	Uncharacterized protein	193547	194773	+	ATG	TAA	1227	408		
gp332	Uncharacterized protein	194760	194966	+	ATG	TAA	207	68		
gp333	Uncharacterized protein	195050	195304	+	TTG	TAA	255	84		
gp334	Uncharacterized protein	195306	195776	+	ATG	TAA	471	156		
gp335	Heat shock protein	195859	196290	+	ATG	TAA	432	143	*Synechococcus* phage S-SM1 heat shock protein	4,00E-17
gp336	Uncharacterized protein	196405	196707	+	ATG	TAA	303	100		
gp337	Structural protein	196782	197330	+	ATG	TAA	549	182		
gp338	Uncharacterized protein	197350	197634	+	ATG	TAA	285	94		
gp339	Uncharacterized protein	197644	197937	+	ATG	TGA	294	97		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp340	Uncharacterized protein	198005	198232	+	ATG	TAA	228	75		
gp341	Uncharacterized protein	198278	198922	+	TTG	TAA	645	214		
gp342	Structural protein	198963	200693	+	ATG	TAA	1731	576		
gp343	DNA ligase	200721	202091	+	ATG	TAA	1371	456	*Vibrio* phage ValKK3 DNA ligase	8,00E-58
gp344	Uncharacterized protein	202101	202262	+	ATG	TAA	162	53		
gp345	Structural protein	202282	202791	+	ATG	TAA	510	169		
gp346	DNA polymerase III	202888	203661	+	ATG	TAA	774	257	*Escherichia* phage 121Q DNA polymerase III	2,00E-29
gp347	Uncharacterized protein	203840	203962	+	ATG	TAA	123	40		
gp348	Uncharacterized protein	203985	204092	+	ATG	TAG	108	35		
gp349	Structural protein	204152	204823	+	ATG	TAA	672	223		
gp350	Uncharacterized protein	205033	205263	-	ATG	TAG	231	76		
gp351	Uncharacterized protein	205351	205608	+	ATG	TAA	258	85		
gp352	Uncharacterized protein	205733	205990	+	ATG	TGA	258	85		
gp353	Uncharacterized protein	205987	206601	+	ATG	TAG	615	204	*Pseudomonas* phage phiPsa374 hypothetical protein	4,00E-13
gp354	Uncharacterized protein	206763	207020	+	ATG	TAA	258	85		
gp355	Uncharacterized protein	207020	207496	+	ATG	TAA	477	158		
gp356	MazG pyrophosphatase	207480	207881	+	ATG	TAA	402	133	*Synechococcus* phage ACG-2014f pyrophosphatase	1,00E-40
gp357	Uncharacterized protein	207841	208278	+	ATG	TAA	438	145		
gp358	Uncharacterized protein	208326	208595	+	ATG	TAA	270	89		
gp359	Uncharacterized protein	208610	208870	+	ATG	TAG	261	86		
gp360	Uncharacterized protein	208891	209103	+	ATG	TGA	213	70		
gp361	Uncharacterized protein	209100	209435	+	ATG	TAA	336	111		
gp362	Uncharacterized protein	209428	209730	+	ATG	TAA	303	100		
gp363	Structural protein	209753	210253	+	ATG	TAA	501	166		
gp364	Uncharacterized protein	210329	210625	+	ATG	TAA	297	98		
gp365	Uncharacterized protein	210626	211333	+	ATG	TAA	708	235		
gp366	DNA polymerase III alpha	211317	211826	+	ATG	TAA	510	169	*Cronobacter* phage vB_CsaM_GAP32 DNA polymerase III alpha	6,00E-28
gp367	Uncharacterized protein	211839	212180	+	ATG	TAA	342	113	*Escherichia* phage 121Q hypothetical protein PBI_121Q_57	1,00E-08
gp368	Uncharacterized protein	212182	212508	+	ATG	TAA	327	108	*Escherichia* phage 121Q chaperonin	2,00E-15
gp369	Clamp loader subunit DNA	212522	213451	+	ATG	TAA	930	309	uncultured Mediterranean phage uvMED clamp loader subunit	5,00E-63
gp370	Structural protein	213453	213815	+	ATG	TAA	363	120		
gp371	Uncharacterized protein	213805	214401	+	ATG	TAA	597	198	*Stenotrophomonas* phage vB_SmaS-DLP_6 hypothetical protein	1,00E-23
gp372	Uncharacterized protein	214484	215257	-	TTG	TAA	774	257		
gp373	Structural protein	215434	215955	+	ATG	TAA	522	173	*Escherichia* phage 121Q structural protein	5,00E-16
gp374	Uncharacterized protein	215985	216686	+	ATG	TAA	702	233	*Escherichia* phage 121Q structural protein	7,00E-17

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp375	Uncharacterized protein	216725	217414	+	ATG	TAA	690	229		
gp376	Uncharacterized protein	217424	219643	+	ATG	TAA	2220	739		
gp377	Structural protein	219761	220195	+	ATG	TAA	435	144	*Cronobacter* phage vB_CsaM_GAP32 hypothetical protein	7,00E-09
gp378	Structural protein	220209	226919	+	ATG	TAA	6711	2236	*Ralstonia* phage RSL1 unnamed protein product	1,00E-21
gp379	Uncharacterized protein	226921	227307	+	ATG	TAA	387	128		
gp380	Structural protein	227423	230443	+	ATG	TAA	3021	1006		
gp381	Structural protein	230447	232420	+	ATG	TAA	1974	657		
gp382	Structural protein	232463	233098	+	ATG	TAA	636	211		
gp383	Structural protein	233165	235849	+	ATG	TAA	2685	894	uncultured Mediterranean phage uvMED phage-related tail fiber	4,00E-08
gp384	Structural protein	236022	236348	+	ATG	TAA	327	108		
gp385	Uncharacterized protein	236556	237128	+	ATG	TAA	573	190	uncultured Mediterranean phage uvMED hypothetical protein	1,00E-21
gp386	Structural protein	237131	238078	+	ATG	TAA	948	315		
gp387	Structural protein	238081	239991	+	ATG	TAA	1911	636	*Pseudomonas* phage PaBG putative adhesin	1,00E-24
gp388	Structural protein	240072	240347	+	ATG	TAA	276	91		
gp389	DNA polymerase	240401	242812	+	ATG	TAA	2412	803	*Cronobacter* phage vB_CsaM_GAP32 DNA polymerase	1,00E-
gp390	Hef-like homing	242805	244487	+	ATG	TAA	1683	560	*Yersinia* phage vB_YenM_TG1 hypothetical protein	8,00E-77
gp391	DNA polymerase	244488	245627	+	ATG	TAA	1140	379	*Cronobacter* phage vB_CsaM_GAP32 DNA polymerase	7,00E-53
gp392	Structural protein	245708	246454	+	ATG	TAA	747	248		
gp393	Uncharacterized protein	246458	246718	+	ATG	TAA	261	86		
gp394	Uncharacterized protein	246791	248413	+	ATG	TAA	1623	540		
gp395	Uncharacterized protein	248413	249474	+	ATG	TAA	1062	353		
gp396	Uncharacterized protein	249559	249936	+	ATG	TAA	378	125		
gp397	Uncharacterized protein	249951	250211	+	ATG	TAA	261	86		
gp398	Uncharacterized protein	250212	250730	+	ATG	TAA	519	172		
gp399	Uncharacterized protein	250742	251176	+	ATG	TAA	435	144		
gp400	Radical SAM domain-	251247	252119	+	ATG	TAA	873	290	*Vibrio* phage ValKK3 radical SAM domain-containing protein	2,00E-52
gp401	Uncharacterized protein	252124	252543	+	ATG	TGA	420	139		
gp402	Uncharacterized protein	252540	253151	+	ATG	TAA	612	203	*Escherichia* phage PBECO 4 hypothetical protein	7,00E-19
gp403	RNaseH ribonuclease	253151	254137	+	ATG	TAA	987	328	*Cronobacter* phage vB_CsaM_GAP32 RNaseH ribonuclease	1,00E-96
gp404	Uncharacterized protein	254140	254598	+	ATG	TAA	459	152		
gp405	Uncharacterized protein	254610	254900	+	ATG	TAA	291	96		
gp406	Uncharacterized protein	254903	255448	+	ATG	TAA	546	181		
gp407	Uncharacterized protein	255451	255651	+	ATG	TAG	201	66		
gp408	Uncharacterized protein	255728	256021	+	ATG	TAA	294	97		
gp409	Uncharacterized protein	256008	256358	+	ATG	TGA	351	116		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp410	Uncharacterized protein	256413	256787	+	ATG	TAA	375	124		
gp411	Uncharacterized protein	256798	257106	+	ATG	TAA	309	102		
gp412	Uncharacterized protein	257172	257681	+	ATG	TAA	510	169		
gp413	Structural protein	257884	258999	+	ATG	TAA	1116	371		
gp414	Uncharacterized protein	259069	259524	+	ATG	TAG	456	151		
gp415	Uncharacterized protein	259526	259984	+	ATG	TGA	459	152		
gp416	Uncharacterized protein	259981	260334	+	ATG	TAA	354	117		
gp417	Homing endonuclease	260334	261830	+	ATG	TAG	1497	498	*Enterobacteria* phage vB_KleM-RaK2 putative homing	3,00E-76
gp418	HNH endonuclease	261886	263070	+	ATG	TAA	1185	394	*Escherichia* phage 121Q Zn-dependent hydrolase	2,00E-86
gp419	Uncharacterized protein	263352	263594	+	ATG	TGA	243	80		
gp420	Uncharacterized protein	263687	264007	-	ATG	TAA	321	106		
gp421	Uncharacterized protein	264174	264413	-	ATG	TAA	240	79		
gp422	Uncharacterized protein	264415	265155	-	ATG	TAA	741	246	*Escherichia* phage 121Q hypothetical protein PBI_121Q_566	1,00E-20
gp423	Uncharacterized protein	265173	265445	-	ATG	TAA	273	90		
gp424	Uncharacterized protein	265438	266040	-	ATG	TAA	603	200		
gp425	Uncharacterized protein	266068	266310	-	ATG	TAA	243	80		
gp426	Structural protein	266297	266545	-	ATG	TAA	249	82		
gp427	Structural protein	266666	267091	-	ATG	TAA	426	141	*Vibrio* phage nt-1 hypothetical protein VPFG_00035	8,00E-07
gp428	Uncharacterized protein	267163	267909	-	ATG	TAA	747	248		
gp429	Structural protein	267902	268861	-	ATG	TAA	960	319		
gp430	Uncharacterized protein	268861	269100	-	ATG	TAA	240	79		
gp431	Uncharacterized protein	269237	269755	-	ATG	TAG	519	172		
gp432	Uncharacterized protein	270165	270266	+	ATG	TAG	102	33		
gp433	Uncharacterized protein	271373	271570	+	ATG	TAA	198	65		
gp434	Uncharacterized protein	271564	271785	+	ATG	TGA	222	73		
gp435	Uncharacterized protein	271782	272000	+	ATG	TGA	219	72		
gp436	Uncharacterized protein	271997	272206	+	ATG	TAA	210	69		
gp437	Uncharacterized protein	272209	272406	+	ATG	TAA	198	65	*Vibrio* phage henriette 12B8 hypothetical protein VPDG_00063	1,00E-05
gp438	Uncharacterized protein	272510	273664	+	ATG	TAA	1155	384	*Cronobacte*r phage S13 hypothetical protein S13_001	1,00E-50
gp439	Uncharacterized protein	273708	274265	+	ATG	TAA	558	185		
gp440	Uncharacterized protein	274234	274641	+	ATG	TAA	408	135		
gp441	Uncharacterized protein	274644	275510	+	ATG	TAA	867	288		
gp442	Uncharacterized protein	275510	276412	+	ATG	TAA	903	300		
gp443	Uncharacterized protein	276415	276927	+	ATG	TAA	513	170		
gp444	Uncharacterized protein	276911	277216	+	TTG	TAA	306	101		

Protein no.	Function	Start	Stop	Strand	Start codon	Stop codon	Lenght (nt)	Lenght (aa)	The closest homolog (BLAST P)	E-value
gp445	Uncharacterized protein	277286	277756	+	ATG	TAA	471	156		
gp446	Uncharacterized protein	277716	277955	+	ATG	TAG	240	79		
gp447	Cell wall hydrolase domain	277966	278532	+	ATG	TAA	567	188	*Pseudomonas* phage vB_PaeM_PAO1_Ab03 putative cell wall	5,00E-55
gp448	Uncharacterized protein	278546	278776	+	ATG	TAA	231	76		
gp449	Uncharacterized protein	278757	279056	+	ATG	TAA	300	99		
gp450	Uncharacterized protein	279131	279490	+	ATG	TAA	360	119		
gp451	Uncharacterized protein	279677	280288	+	ATG	TAA	612	203		
gp452	Uncharacterized protein	280282	280890	+	ATG	TGA	609	202		
gp453	Uncharacterized protein	280883	281362	+	ATG	TAA	480	159		
gp454	Uncharacterized protein	281397	281774	+	ATG	TGA	378	125		
gp455	Uncharacterized protein	281771	282397	+	ATG	TAG	627	208		
gp456	Uncharacterized protein	282408	282737	+	ATG	TAA	330	109		
gp457	Uncharacterized protein	282727	284154	+	ATG	TAA	1428	475		
gp458	Uncharacterized protein	284160	284423	+	ATG	TAA	264	87		
gp459	Structural protein	284454	285206	-	ATG	TAA	753	250		
gp460	Structural protein	285187	285876	-	ATG	TAA	690	229		
gp461	Uncharacterized protein	285883	286077	-	ATG	TAG	195	64		
gp462	Structural protein	286163	287143	+	ATG	TAA	981	326		

Table S 6.3: The ESI-MS/MS analysis of denaturated phage PA5oct particles after fractionation on SDSPAGE gel.

Protein no.	Identified function	Band no.	Molecular weight (Da)		Number of identified peptides	Sequence coverage	Identity with other pahges (BLASTP)				
			MS	SDS-PAGE			GAP32	PBECO 4	121Q	KleM-RaK2	K64-1
gp20	Structural protein	15	14 003,90	13090	4	44,53%					
gp100	Structural protein	11	26 989,70	25740	5	28,20%					
gp112	Structural protein	15	8 566,50	7150	2	35,38%					
gp115	Structural protein	13	22 088,90	20020	2	29,12%					
gp204	Structural protein	7	44 484,70	44660	3	21,67%					
gp205	Structural protein	6	50 631,60	52580	12	62,34%					
gp206	Structural protein	5	62 159,40	63910	21	65,58%					
gp207	Structural protein	12	22 521,10	22110	10	69,65%					
gp209	Structural protein	1	437 680,80	424380	107	43,65%					
gp210	Baseplate wedge	2	150 682,40	145090	52	63,76%	28%	27%	27%	32%	32%
gp211	Baseplate protein	15	16 540,30	14300	5	49,23%	37%	-	33%	33%	-
gp214	Baseplate hub subunit and tail lysozyme	6	51 283,90	52470	17	57,65%	31%	30%	30%	32%	32%
gp215	Structural protein	4	79 369,90	74140	8	19,88%	21%	-	21%	22%	22%
gp217	Structural protein	11	24 530,10	23210	8	77,25%	31%	-	29%	31%	-
gp220	Recombination endonuclease subunit*	5	68 787,50	64240	10	25,00%	35%	34%	34%	33%	-
gp221	Structural protein	11	28 585,50	27390	10	44,98%					
gp227	Structural protein	4	86 353,30	87780	27	49,87%					
gp228	Portal vertex protein of head	5	65 729,50	62810	17	48,69%	44%	43%	42%	43%	-
gp229	Structural protein	15	10 165,00	8030	4	67,12%	48%	-	38%	41%	-
gp230	Structural protein	10	30 221,10	29700	8	44,07%					
gp231	Prohead core scaffolding protein and protease	12	23 825,40	23100	6	40,00%	48%	-	53%	44%	-
gp233	Precursor of major head subunit	7	43 001,30	43450	18	97,47%	68%	-	66%	66%	-
gp238	Phage tail fiber protein H	3	106 898,80	110880	23	50,20%					
gp240	Tail sheath stabilizer and completion protein	6	52 038,00	50160	13	38,60%	28%	-	32%	36%	36%
gp241	Structural protein	8	41 824,70	39930	2	6,34%	34%	30%	30%	32%	32%
gp243	Tail fiber protein	6	57 453,40	56210	19	48,34%					
gp244	Structural protein	5	60 641,20	58520	20	64,47%					
gp245	Structural protein	4	86 840,50	86570	18	35,07%					
gp247	Structural protein	15	16 690,10	14960	4	43,38%					
gp248	Structural protein	15	12 844,40	11550	2	42,86%					

Protein no.	Identified function	Band no.	Molecular weight (Da)		Number of identified peptides	Sequence coverage	Identity with other pahges (BLASTP)				
			MS	SDS-PAGE			GAP32	PBECO 4	121Q	KleM-RaK2	K64-1
gp250	Glycosyl hydrolase, lysozyme-like protein	14	18 902,50	18810	5	45,03%					
gp252	Structural protein	15	9 114,50	7700	2	25,71%					
gp257	Structural protein	13	21 738,70	20570	5	39,57%					
gp264	RIIB protector from prophage-induced early lysis*	10	29 135,10	27170	10	53,04%					
gp270	Neck protein	10	29 133,00	27500	9	62,40%	42%	40%	40%	40%	-
gp271	Structural protein	11	26 823,60	23430	5	32,86%	45%	42%	41%	43%	-
gp272	Tail sheath monomer	3	95 769,60	96690	37	80,55%	52%	-	51%	50%	50%
gp274	Structural protein	10	29 474,00	22660	9	82,04%	44%	36%	44%	40%	-
gp275	Structural protein	5	59 138,80	54120	16	52,24%					
gp276	Structural protein	15	17 211,80	16610	4	50,33%					
gp279	Structural protein	12	22 977,90	22000	13	98,00%					
gp280	Structural protein	6	55 894,10	55110	10	25,95%					
gp282	Structural protein	5	59 957,40	54560	20	55,04%					
gp283	Structural protein	6	57 679,30	54560	11	28,02%	52%	-	38%	40%	40%
gp284	Structural protein	9	37 579,30	36080	8	40,55%	42%	-	45%	43%	-
gp287	RecA-like protein	6	47 297,00	46530	2	7,56%	59%	-	58%	56%	-
gp298	Structural protein	9	35 821,40	31790	4	23,18%					
gp302	Structural protein	12	23 295,90	22330	7	48,77%					
gp303	Structural protein	7	45 019,60	43890	13	47,37%					
gp306	Structural protein	15	9 395,90	8910	3	35,80%					
gp307	Structural protein	11	28 071,10	26070	6	56,96%					
gp308	Structural protein	14	18 165,80	15730	6	55,94%					
gp309	Structural protein	15	11 193,40	11000	4	87,00%					
gp310	Structural protein	15	17 023,00	16940	4	48,70%					
gp311	Structural protein	11	28 225,90	20240	2	11,41%					
gp313	Structural protein	7	44 353,50	43450	9	27,34%					
gp314	Structural protein	15	11 322,20	11440	2	22,12%					
gp315	Structural protein	15	13 692,40	12210	2	32%					
gp317	Structural protein	7	45 172,10	43890	11	44,86%					
gp319	Structural protein	4	88 245,70	83710	3	8,54%					

Protein no.	Identified function	Band no.	Molecular weight (Da)		Number of identified peptides	Sequence coverage	Identity with other pahges (BLASTP)				
			MS	SDS-PAGE			GAP32	PBECO 4	121Q	KleM-RaK2	K64-1
gp320	Structural protein	7	45 858,00	41690	7	20,84%					
gp321	Putative baseplate hub sub and tail lysozyme, muramoyl-pentapeptidase	8	37 776,60	37400	4	15,88%					
gp327	Structural protein	9	34 524,70	32780	9	37,25%					
gp329	GTP cyclohydrolase*	11	26 338,70	25190	13	79,91%					
gp330	Structural protein	13	21 589,60	18920	3	26,16%					
gp337	Structural protein	12	23 119,90	20020	3	31,32%					
gp342	Structural protein	5	63 016,60	63360	2	2,95%					
gp343	DNA ligase*	6	52 246,10	50160	2	7,02%	34%	33%	33%	-	-
gp345	Structural protein	13	19 181,60	18590	4	39%					
gp349	Structural protein	11	26 446,00	24530	5	33,18%					
gp356	MazG pyrophosphatese*	15	15 569,30	14630	1	18,04%					
gp363	Structural protein	13	21 637,60	18260	7	57,23%					
gp370	Structural protein	15	14 739,40	13200	3	30,00%					
gp373	Structural protein	13	20 812,50	19030	9	60,69%	29%	34%	34%	30%	-
gp377	Structural protein	14	18 313,30	15840	5	59,72%	38%	-	33%	31%	-
gp378	Structural protein	1	237 196,80	245960	31	29,20%					
gp380	Structural protein	3	106 920,50	110660	8	16,89%					
gp381	Structural protein	4	71 076,70	72270	8	17,20%					
gp382	Structural protein	12	22 605,90	23210	4	29,38%					
gp383	Structural protein	4	91 998,00	98340	16	40,60%					
gp384	Structural protein	15	13 740,70	11880	2	39,81%					
gp386	Structural protein	9	34 454,70	34650	8	43,49%					
gp387	Structural protein	4	70 630,00	69960	15	37,58%					
gp388	Structural protein	15	11 452,20	10010	4	65,93%					
gp392	Structural protein	11	27 239,60	27280	5	39,64%					
gp403	RNaseH ribonuclease*	8	37 919,60	36080	5	24,39%	46%	44%	45%	45%	-
gp413	Structural protein	7	42 181,70	40810	7	22,64%					
gp426	Structural protein	15	10 749,90	9020	3	39,02%					
gp427	Structural protein	15	16 830,80	15510	8	81,56%					
gp429	Structural protein	8	37 720,50	35090	8	36,05%					

Protein no.	Identified function	Band no.	Molecular weight (Da)		Number of identified peptides	Sequence coverage	Identity with other pahges (BLASTP)				
			MS	SDS-PAGE			GAP32	PBECO 4	121Q	KleM-RaK2	K64-1
gp459	Structural protein	10	29 986,60	27500	13	77,20%					
gp460	Structural protein	11	27 328,40	25190	12	68,56%					
gp462	Structural protein	9	36 516,80	35860	11	50,92%					

(*) proteins non-associated with phage particles

Table S 6.4: List of jumbo phages and relevant phages that belong to the viral vlusters (VCs).

VC_77		
Accession No.	Name	Taxonomy
NC_019526.1	*Enterobacteria* phage vB_KleM-RaK2	*unclassified Myoviridae*
NC_025447.1	*Escherichia* phage 121Q	*unclassified Myoviridae*
NC_027399.1	*Klebsiella* phage K64-1	*unclassified Myoviridae*
NC_027364.1	*Escherichia* phage PBECO 4	*unclassified Myoviridae*
PA5oct		
NC_019401.1	*Cronobacter* phage vB_CsaM_GAP32	*unclassified Myoviridae*
VC_87		
Accession No.	Name	Taxonomy
NC_019410.1	*Caulobacter* phage karma	*Phicbkvirus*
NC_019411.1	*Caulobacter* phage swift	*Phicbkvirus*
NC_019405.1	*Caulobacter* phage phiCbK	*Phicbkvirus*
NC_019407.1	*Caulobacter* phage magneto	*Phicbkvirus*
NC_019406.1	*Caulobacter* phage CcrColossus	*unclassified Siphoviridae*
NC_019408.1	*Caulobacter* phage rogue	*Phicbkvirus*
VC_172		
Accession No.	Name	Taxonomy
NC_022096.1	Pseudomonas phage PaBG	unclassified Myoviridae
NC_017972.1	Pseudomonas phage Lu11	unclassified Myoviridae
NC_010811.2	Ralstonia phage RSL1	unclassified Myoviridae
VC_44		
Accession No.	Name	Taxonomy
NC_028899.1	*Ralstonia* phage RSF1	*unclassified Myoviridae*
NC_004629.1	*Pseudomonas* phage φKZ	*Phikzvirus*
NC_028999.1	*Pseudomonas* phage PhiPA3	*Phikzvirus*
NC_007623.1	*Pseudomonas* phage EL	*Phikzvirus*
NC_016571.1	*Pseudomonas* phage OBP	*unclassified Myoviridae*
NC_010821.1	*Pseudomonas* phage 201phi2-1	*Phikzvirus*
NC_017975.1	*Vibrio* phage JM-2012	*unclassified Myoviridae*
NC_023610.1	*Erwinia* phage PhiEaH1	*unclassified Myoviridae*
NC_023557.1	*Erwinia* phage Ea35-70	*unclassified Myoviridae*
NC_016163.1	*Yersinia* phage phiR1-37	*unclassified Myoviridae*
NC_028950.1	*Ralstonia* phage RSL2	*unclassified Myoviridae*

VC_150		
Accession No.	Name	Taxonomy
NC_027402.1	*Salmonella* phage SPN3US	*unclassified Myoviridae*
NC_021531.1	*Cronobacter* phage CR5	*unclassified Myoviridae*
NC_019929.1	*Erwinia* phage phiEaH2	*unclassified Siphoviridae*
VC_43		
Accession No.	Name	Taxonomy
NC_023007.1	*Bacillus* phage vB_BanS-Tsamsa	*unclassified Siphoviridae*
NC_007581.1	*Clostridium* phage c-st	*unclassified Myoviridae*
NC_029073.1	*Geobacillus* virus E3	*unclassified Siphoviridae*
NC_023574.1	*Lactococcus* phage phiL47	*unclassified Siphoviridae*
NC_001884.1	*Bacillus* phage SPbeta	*Spbetavirus*
NC_027341.1	*Lactococcus* phage WRP3	*unclassified Siphoviridae*
NC_029048.1	*Clostridium* phage phiCD211	*unclassified Myoviridae*
NC_028749.1	*Brevibacillus* phage Sundance	*unclassified Siphoviridae*
NC_015263.1	*Lactococcus* phage 949	*unclassified Siphoviridae*
NC_029119.1	*Staphylococcus* phage SPbeta-like	*Spbetavirus*
NC_023719.1	*Bacillus* phage G	*unclassified Myoviridae*
VC_184		
Accession No.	Name	Taxonomy
NC_028805.1	*Brevibacillus* phage Jenst	*unclassified Siphoviridae*
NC_009760.1	*Bacillus* phage 0305φ8-36	*unclassified Myoviridae*

Table S 6.5: List of jumbo phages and relevant phages that belong to the viral vlusters (VCs).

VC_2		
Accession No.	Name	Taxonomy
NC_015251.1	*Aeromonas* phage 65	*T4virus*
NC_005260.1	*Aeromonas* phage Aeh1	*T4virus*
NC_019538.1	*Aeromonas* phage CC2	*T4virus*
NC_014636.1	*Aeromonas* phage phiAS5	*T4virus*
NC_023688.1	*Aeromonas* phage PX29	*T4virus*
NC_000866.4	*Enterobacteria* phage T4	*T4virus*
NC_004928.1	*Enterobacteria* phage RB69	*T4virus*
NC_005066.1	*Enterobacteria* phage RB49	*T4virus*
NC_005135.1	*Aeromonas* phage 44RR2.8t	*T4virus*
NC_007022.1	*Aeromonas* phage 31	*T4virus*
NC_007023.1	*Enterobacteria* phage RB43	*T4virus*
NC_008208.1	*Aeromonas* phage 25	*T4virus*
NC_008515.1	*Enterobacteria* phage RB32	*T4virus*
NC_009821.1	*Enterobacteria* phage Phi1	*T4virus*
NC_010105.1	*Enterobacteria* phage JS98	*T4virus*
NC_012635.1	*Enterobacteria* phage RB51	*T4virus*
NC_012638.1	*Enterobacteria* phage RB14	*T4virus*
NC_012740.1	*Enterobacteria* phage JSE	*T4virus*
NC_012741.1	*Enterobacteria* phage JS10	*T4virus*
NC_014036.1	*Klebsiella* phage KP15	*T4virus*
NC_014260.1	*Enterobacteria* phage IME08	*T4virus*
NC_014467.1	*Enterobacteria* phage RB16	*T4virus*
NC_014595.1	*Shigella* phage SP18	*T4virus*
NC_014635.1	*Aeromonas* phage phiAS4	*T4virus*
NC_014660.1	*Acinetobacter* phage Ac42	*T4virus*
NC_014661.1	*Acinetobacter* phage Acj61	*T4virus*
NC_014662.1	*Enterobacteria* phage CC31	*T4virus*
NC_014663.1	*Acinetobacter* phage Acj9	*T4virus*
NC_014792.1	*Enterobacter*ia phage vB_EcoM-VR7	*T4virus*
NC_015250.1	*Acinetobacter* phage 133	*T4virus*
NC_015457.1	*Shigella* phage Shfl2	*T4virus*
NC_015464.1	*Campylobacter* phage NCTC12673	*Cp8unavirus*
NC_016562.1	*Campylobacter* phage CPX	*Cp8unavirus*
NC_018087.3	*Acinetobacter* phage ZZ1	*T4virus*

VC_8		
Accession No.	Name	Taxonomy
NC_006820.1	*Synechococcus* phage S-PM2	*T4virus*
NC_006884.2	*Prochlorococcus* phage P-SSM4	*T4virus*
NC_008296.2	*Synechococcus* phage syn9	*T4virus*
NC_013085.1	*Synechococcus* phage S-RSM4	*T4virus*
NC_015279.1	*Synechococcus* phage S-SM2	*unclassified myoviridae*
NC_015280.1	*Prochlorococcus* phage P-HM1	*unclassified myoviridae*
NC_015281.1	*Synechococcus* phage S-ShM2	*unclassified myoviridae*
NC_015282.1	*Synechococcus* phage S-SM1	*unclassified myoviridae*
NC_015283.1	*Prochlorococcus* phage P-RSM4	*unclassified myoviridae*
NC_015284.1	*Prochlorococcus* phage P-HM2	*unclassified myoviridae*
NC_015285.1	*Prochlorococcus* phage Syn33	*unclassified myoviridae*
NC_015286.1	*Synechococcus* phage Syn19	*unclassified myoviridae*
NC_015288.1	*Prochlorococcus* phage Syn1	*unclassified myoviridae*
NC_015289.1	*Synechococcus* phage S-SSM5	*unclassified myoviridae*
NC_015290.1	*Prochlorococcus* phage P-SSM7	*unclassified myoviridae*
NC_015569.1	*Synechococcus* phage S-CRM01	*unclassified myoviridae*
NC_019443.1	*Synechococcus* phage metaG-MbCM1	*T4virus*
NC_019444.1	*Synechococcus* phage ACG-2014c	*T4virus*
NC_020484.1	*Pelagibacter* phage HTVC008M	*unclassified Myoviridae*
NC_020486.1	*Synechococcus* phage S-RIM8	*unclassified Myoviridae*
NC_020837.1	*Synechococcus* phage S-CAM1	*unclassified Myoviridae*
NC_020845.1	*Prochlorococcus* phage MED4-213	*unclassified Myoviridae*
NC_020855.1	Cyanophage P-RSM6	*unclassified Myoviridae*
NC_020859.1	*Synechococcus* phage S-RIM2	*unclassified Myoviridae*
NC_020875.1	*Synechococcus* phage S-SSM4	*unclassified Myoviridae*
NC_021071.1	*Cyanophage* P-RSM1	*unclassified Myoviridae*
NC_021072.1	Cyanophage Syn30	*unclassified dsDNA phages*
NC_021530.1	*Synechococcus* phage S-CAM8	*unclassified Myoviridae*
NC_021536.1	*Synechococcus* phage S-IOM18	*unclassified Myoviridae*
NC_021559.1	*Prochlorococcus* phage P-SSM3	*unclassified Myoviridae*
NC_023584.1	*Synechococcus* phage S-MbCM100	*T4virus*
NC_023587.1	*Synechococcus* phage ACG-2014h	*T4virus*
NC_025422.1	Caulobacter phage Cr30	*unclassified Myoviridae*
NC_026923.1	*Synechococcus* phage ACG-2014d	*T4virus*

NC_018855.1	*Enterobacteria* phage HX01	*T4virus*
NC_018861.1	*Campylobacter* phage CP30A	*unclassified Myoviridae*
VC_2		
NC_019398.1	*Cronobacter* phage vB_CsaM_GAP161	*T4virus*
NC_019399.1	*Enterobacteria* phage vB_EcoM_ACG-C40	*T4virus*
NC_019500.1	*Enterobacteria* phage Bp7	*T4virus*
NC_019503.1	*Enterobacteria* phage ime09	*T4virus*
NC_019505.1	*Escherichia* phage wV7	*T4virus*
NC_019543.1	*Aeromonas* phage Aes508	*T4virus*
NC_019909.1	*Yersinia* phage phiR1-RT	*unclassified Myoviridae*
NC_020080.1	*Klebsiella* phage KP27	*T4virus*
NC_020416.1	*Salmonella* phage S16	*T4virus*
NC_020879.1	*Aeromonas* phage Aes012	*T4virus*
NC_021073.1	*Vibrio* phage henriette 12B8	*unclassified dsDNA phages*
NC_021344.2	*Escherichia* phage Lw1	*T4virus*
NC_023561.1	*Enterobacter* phage PG7	*unclassified Myoviridae*
NC_024121.1	*Serratia* phage PS2	*T4virus*
NC_024124.2	*Escherichia* phage vB_EcoM_JS09	*T4virus*
NC_024125.2	*Escherichia* phage e11/2	*T4virus*
NC_024794.1	*Escherichia* phage vB_EcoM_PhAPEC2	*T4virus*
NC_025414.1	*Citrobacter* phage Miller	*T4virus*
NC_025419.1	*Escherichia* phage T4	*T4virus*
NC_025425.1	*Enterobacteria* phage GEC-3S	*T4virus*
NC_025429.1	*Rhizobium* phage vB_RleM_P10VF	*unclassified Myoviridae*
NC_025437.1	*Shigella* phage Shf125875	*unclassified Myoviridae*
NC_025448.1	*Escherichia* phage T4	*T4virus*
NC_025449.1	*Escherichia* phage ECML-134	*T4virus*
NC_025829.1	*Shigella* phage pSs-1	*unclassified Myoviridae*
NC_026607.2	*Salmonella* phage STP4-a	*T4virus*
NC_027204.1	*Sinorhizobium* phage phiM19	*T4virus*
NC_027331.1	*Citrobacter* phage Moon	*unclassified Myoviridae*
NC_027344.1	*Salmonella* phage STML-198	*unclassified Myoviridae*
NC_027349.1	*Escherichia* phage HY01	*unclassified Myoviridae*
NC_027353.1	*Yersinia* phage phiD1	*unclassified Myoviridae*
NC_027404.1	*Yersinia* phage PST	*T4virus*
NC_027979.1	*Escherichia* phage T4	*T4virus*
NC_026924.1	*Synechococcus* phage ACG-2014g	*T4virus*
NC_026926.1	*Synechococcus* phage ACG-2014j	*T4virus*
VC_8		
NC_026928.1	*Synechococcus* phage ACG-2014e	*T4virus*
NC_027130.1	*Synechococcus* phage ACG-2014b	*T4virus*
NC_027132.1	*Synechococcus* phage ACG-2014i	*T4virus*
NC_028663.1	Cyanophage P-TIM40	*unclassified Myoviridae*
NC_028955.1	*Prochlorococcus* phage P-TIM68	*unclassified Myoviridae*
VC_18		
Accession No.	Name	Taxonomy
NC_013693.1	*Shigella* phage Ag3	*Viunavirus*
NC_013697.1	*Deftia* phage phiW-14	*unclassified myoviridae*
NC_015296.1	*Salmonella* phage ViI	*unclassified myoviridae*
NC_016073.1	*Salmonella* phage SFP10	*Viunavirus*
NC_016570.1	*Escherichia* phage Cba120	*Viunavirus*
NC_019452.1	*Escherichia* phage PhaxI	*Viunavirus*
NC_019530.1	*Salmonella* phage PhiSH19	*Viunavirus*
NC_019910.1	*Salmonella* phage SKML-39	*unclassified Myoviridae*
NC_019925.1	*Dickeya* phage Limestone	*Viunavirus*
NC_020083.1	*Serratia* phage phiMAM1	*unclassified Myoviridae*
NC_022343.1	*Klebsiella* phage 0507-KN2-1	*unclassified Myoviridae*
NC_022768.1	*Salmonella* phage Maynard	*unclassified Myoviridae*
NC_022772.1	*Salmonella* phage Marshall	*unclassified Myoviridae*
NC_023856.1	*Salmonella* phage vB_SalM_SJ2	*unclassified Myoviridae*
NC_024122.1	*Salmonella* phage vB_SalM_SJ3	*unclassified Myoviridae*
NC_025446.1	*Escherichia* phage ECML-4	*Viunavirus*
NC_025452.1	*Dickeya* phage RC-2014	*unclassified Myoviridae*
NC_027119.1	*Salmonella* phage Det7	*unclassified Myoviridae*
NC_027340.1	*Erwinia* phage phiEa2809	*unclassified Myoviridae*
NC_029042.1	*Salmonella* phage 38	*unclassified Myoviridae*
VC_111		
Accession No.	Name	Taxonomy
NC_019507.1	*Campylobacter* phage CP21	*Cp220virus*
NC_027996.1	*Campylobacter* phage CPt10	*Cp220virus*
NC_027997.1	*Campylobacter* phage CP220	*Cp220virus*
VC_150		

NC_027983.1	*Enterobacteria* phage AR1	*T4virus*
NC_028448.1	*Escherichia* phage slur08	*T4virus*
NC_028676.1	*Sinorhizobium* phage phiM9	*unclassified Myoviridae*
NC_028683.1	*Edwardsiella* phage PEi20	*unclassified Myoviridae*
VC_2		
NC_028686.1	*Klebsiella* phage JD18	*unclassified Myoviridae*
NC_028762.1	*Proteus* phage vB_PmiM_Pm5461	*T4virus*
NC_028773.1	*Cronobacter* phage S13	*unclassified Myoviridae*
NC_028780.1	*Escherichia* phage slur07	*T4virus*
NC_028820.1	*Yersinia* phage vB_YenM_TG1	*unclassified Myoviridae*
NC_028829.1	*Vibrio phage* ValKK3	*unclassified Caudovirales*
NC_028847.1	*Enterobacteria* phage QL01	*T4virus*
NC_028881.1	*Enterobacteria* phage vB_EcoM_VR5	*T4virus*
NC_028894.1	*Enterobacteria* phage vB_EcoM_VR20	*T4virus*
NC_028925.1	*Enterobacteria* phage vB_EcoM_VR25	*T4virus*
NC_028927.1	*Escherichia* phage slur02	*T4virus*
NC_028940.1	*Pectobacterium* bacteriophage PM2	*unclassified Myoviridae*
NC_028957.1	*Enterobacteria* phage vB_EcoM_VR26	*T4virus*
NC_029000.1	*Stenotrophomonas* phage IME13	*T4virus*
NC_029013.1	*Citrobacter* phage IME-CF2	*T4virus*
NC_029091.1	*Escherichia* phage APCEc01	*unclassified Myoviridae*
NC_006883.2	*Prochlorococcus* phage P-SSM2	*T4virus*
NC_028945.1	*Sinorhizobium* phage phiN3	*T4virus*
NC_019521.1	*Sphingomonas* phage PAU	*unclassified Myoviridae*
NC_026927.1	*Synechococcus* phage ACG-2014f	*T4virus*
NC_020851.1	*Synechococcus* phage S-SKS1	*unclassified Siphoviridae*
NC_015287.1	*Synechococcus* phage S-SSM7	*unclassified Myoviridae*
NC_021529.2	*Vibio* phage nt-1	*SchizoT4virus*
NC_005083.2	*Vibrio* phage KVP40	*SchizoT4virus*
NC_023568.1	*Vibrio* phage VH7D	*SchizoT4virus*
VC_44		
Accession No.	Name	Taxonomy
NC_028899.1	*Ralstonia* phage RSF1	*unclassified Myoviridae*
NC_004629.1	*Pseudomonas* phage φKZ	*Phikzvirus*

Accession No.	Name	Taxonomy
NC_027402.1	*Salmonella* phage SPN3US	*unclassified Myoviridae*
NC_021531.1	*Cronobacter* phage CR5	*unclassified Myoviridae*
NC_019929.1	*Erwinia* phage phiEaH2	*unclassified Siphoviridae*
VC_43		
Accession No.	Name	Taxonomy
NC_023007.1	*Bacillus* phage vB_BanS-Tsamsa	*unclassified Siphoviridae*
NC_007581.1	*Clostridium* phage c-st	*unclassified Myoviridae*
NC_029073.1	*Geobacillus* virus E3	*unclassified Siphoviridae*
NC_023574.1	*Lactococcus* phage phiL47	*unclassified Siphoviridae*
NC_001884.1	*Bacillus* phage SPbeta	*Spbetavirus*
NC_027341.1	*Lactococcus* phage WRP3	*unclassified Siphoviridae*
NC_029048.1	*Clostridium* phage phiCD211	*unclassified Myoviridae*
NC_028749.1	*Brevibacillus* phage Sundance	*unclassified Siphoviridae*
NC_015263.1	*Lactococcus* phage 949	*unclassified Siphoviridae*
NC_029119.1	*Staphylococcus* phage SPbeta-like	*Spbetavirus*
NC_023719.1	*Bacillus* phage G	*unclassified Myoviridae*
VC_184		
Accession No.	Name	Taxonomy
NC_028805.1	*Brevibacillus* phage Jenst	*unclassified Siphoviridae*
NC_009760.1	*Bacillus* phage 0305φ8-36	*unclassified Myoviridae*
VC_77		
Accession No.	Name	Taxonomy
NC_019526.1	*Enterobacteria* phage vB_KleM-RaK2	*unclassified Myoviridae*
NC_025447.1	*Escherichia* phage 121Q	*unclassified Myoviridae*
NC_027399.1	*Klebsiella* phage K64-1	*unclassified Myoviridae*
NC_027364.1	*Escherichia* phage PBECO 4	*unclassified Myoviridae*
	Pseudomonas PA5oct	
NC_019401.1	*Cronobacter* phage vB_CsaM_GAP32	*unclassified Myoviridae*
VC_87		
Accession No.	Name	Taxonomy
NC_019410.1	*Caulobacter* phage karma	*Phicbkvirus*
NC_019411.1	*Caulobacter* phage swift	*Phicbkvirus*

NC_028999.1	*Pseudomonas* phage PhiPA3	*Phikzvirus*
NC_007623.1	*Pseudomonas* phage EL	*Phikzvirus*
NC_016571.1	*Pseudomonas* phage OBP	*unclassified Myoviridae*
NC_010821.1	*Pseudomonas* phage 201phi2-1	*Phikzvirus*
NC_017975.1	*Vibrio* phage JM-2012	*unclassified Myoviridae*
NC_023610.1	*Erwinia* phage PhiEaH1	*unclassified Myoviridae*
NC_023557.1	*Erwinia* phage Ea35-70	*unclassified Myoviridae*
VC_44		
NC_016163.1	*Yersinia* phage phiR1-37	*unclassified Myoviridae*
NC_028950.1	*Ralstonia* phage RSL2	*unclassified Myoviridae*

NC_019405.1	*Caulobacter* phage phiCbK	*Phicbkvirus*
NC_019407.1	*Caulobacter* phage magneto	*Phicbkvirus*
NC_019406.1	*Caulobacter* phage CcrColossus	*unclassified Siphoviridae*
NC_019408.1	*Caulobacter* phage rogue	*Phicbkvirus*
VC_172		
Accession No.	Name	Taxonomy
NC_022096.1	*Pseudomonas* phage PaBG	*unclassified Myoviridae*
VC_172		
NC_017972.1	*Pseudomonas* phage Lu11	*unclassified Myoviridae*
NC_010811.2	*Ralstonia* phage RSL1	*unclassified Myoviridae*

Supplementary data- CHAPTER 7

LKA1 gp49

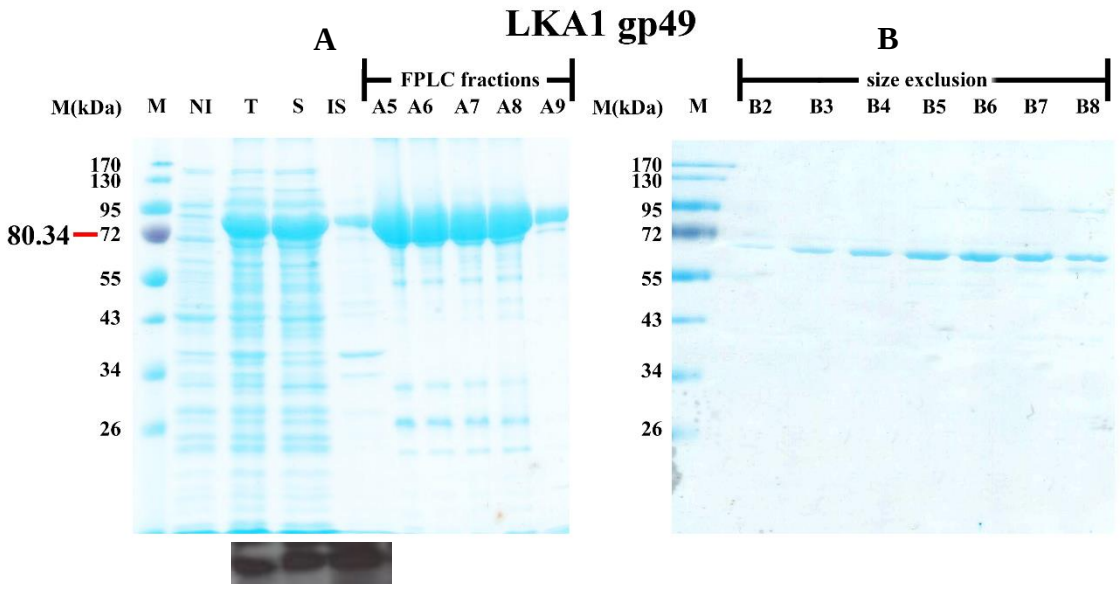


Figure. S 7.1: Expression and purification analysis in optimal conditions of LKA1 gp49. (A) 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Prestained Marker; 2) recombinant protein expression in (NI) not induced, (T) total, (S) soluble and (IS) insoluble fractions processed from 10 ml sample (10µl/well); 3) affinity chromatography fractions from FPLC purification, fractions A5 - A9 (10µl/well). Under T, S, IS fractions corresponding bands from Western Blot are presented. Expected protein size (80.34 kDa) is marked on the left side of the gel with a red band. **(B)** 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Prestained Marker; 2) size exclusion chromatography fractions B2 - B8 (10µl/well). The SDS-PAGE gels are stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific).

LKA1 gp49 domain

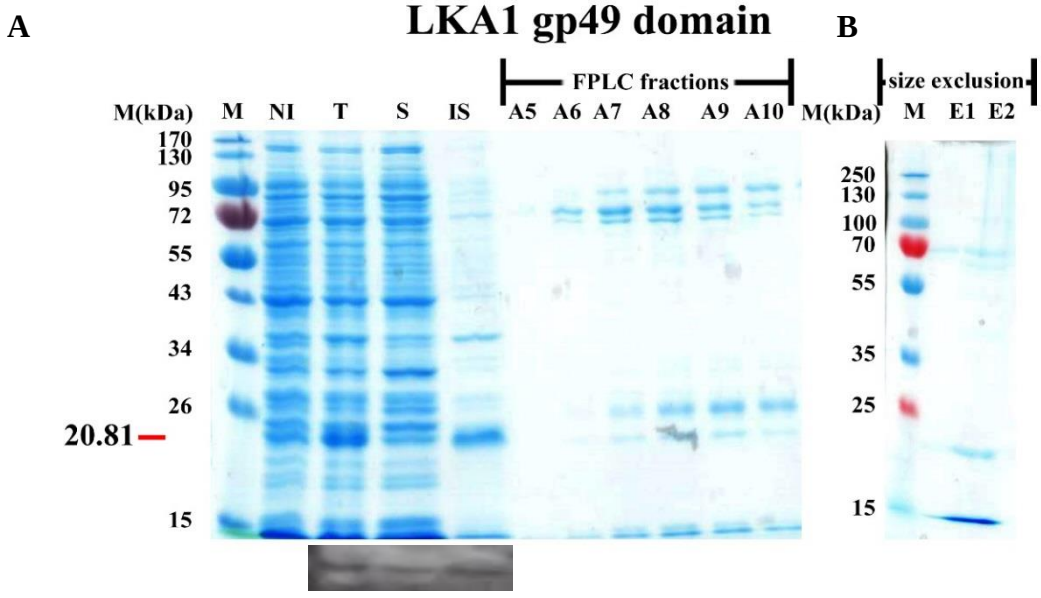


Figure. S 7.17: Expression and purification analysis in optimal conditions of LKA1 gp49 domain. (A) 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Prestained Marker; 2) recombinant protein expression in (NI) not induced, (T) total, (S) soluble and (IS) insoluble fractions processed from 10 ml sample (10µl/well); 3) affinity chromatography fractions from FPLC purification, fractions A5 - A10

(10µl/well). Under T, S, IS fractions corresponding bands from Western Blot are presented. Expected protein size (20.81 kDa) is marked on the left side of the gel with a red band. **(B)** 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Plus Prestained Marker; 2) size exclusion chromatography fractions E1-E2 from protein peak no. 5 (10µl/well). The SDS-PAGE gels are stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific).

LUZ7 gp56 domain

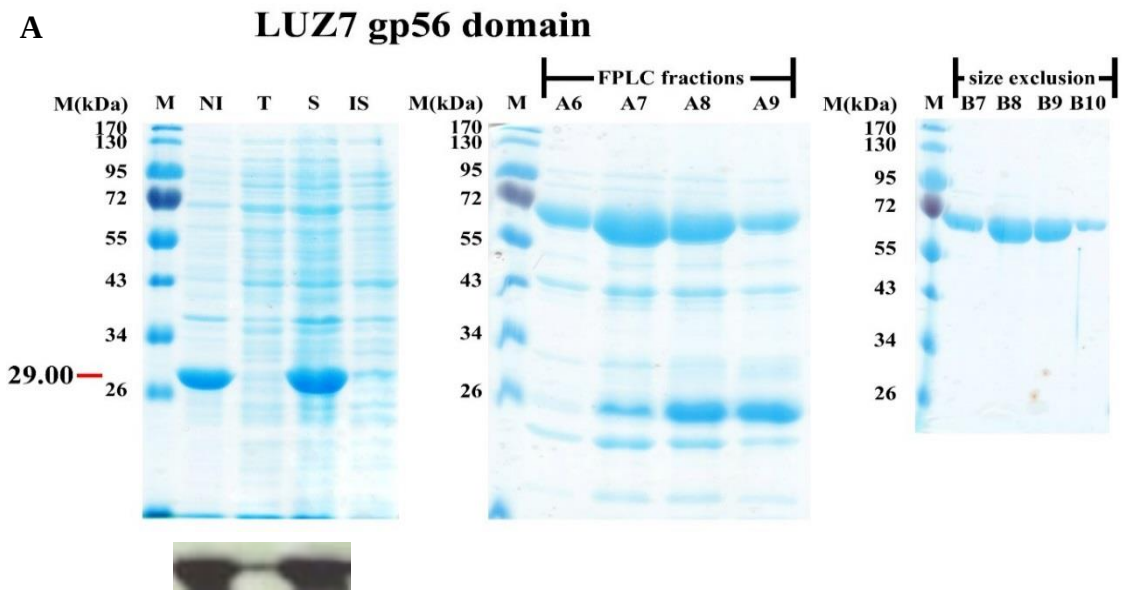


Figure. S 7.18: Expression and purification analysis in optimal conditions of LUZ7 gp56 domain. (A) On the left site is a 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Prestained Marker; 2) recombinant protein expression in (NI) not induced, (T) total, (S) soluble and (IS) insoluble fractions processed from 10 ml sample (10µl/well); 3) affinity chromatography fractions from FPLC purification, fractions A6 - A9 (10µl/well). Under T, S, IS fractions corresponding bands from Western Blot are presented. Expected protein size (29.00 kDa) is marked on the left side of the gel with a red band. On the right side is a 12% acrylamide SDS-PAGE gel with: 1) (M) PageRuler™ Plus Prestained Marker; 2) size exclusion chromatography fractions B7-B10 from protein peak no. 1 (10µl/well). The SDS-PAGE gels are stained with GelCode™ Blue Safe Protein Stain (Thermo Scientific).

Supplementary data- CHAPTER 7

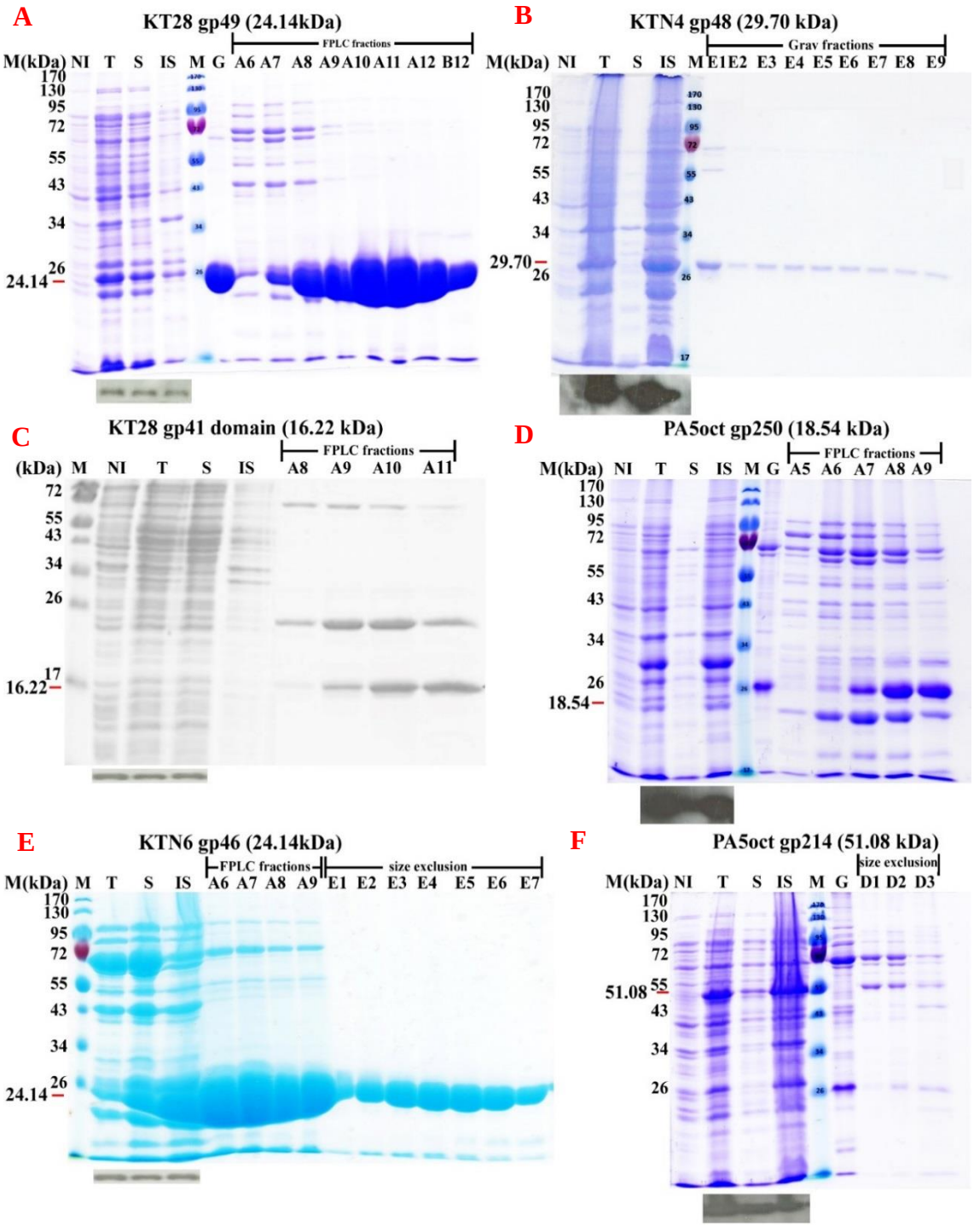


Figure S 8.1: Expression and purification analysis of PG degrading enzymes. (A) KT28 gp49 (B) KTN4 gp48 (C) KT28 gp41, (D) PA5oct gp250, (E) KTN6 gp46, (F) PA5octgp214. Abbreviations: (NI) not induced, (T) total, (S) soluble, (IS) insoluble fractions processed from 10 ml sample;(M) PageRuler™ Prestained Marker; (G) protein purification fractions collected from gravity Ni^{2+}-NTA affinity chromatography; (FPLC) protein purification fractions collected from affinity chromatography with the use of HisTrap™ HP 1 ml columns and Akta FPLC-system; (size exclusion) protein purified with the use of size exclusion chromatography. Each fraction is loaded on a gel in volume of 10µl/well. Under T, S, IS fractions corresponding bands from Western Blot are presented. Expected protein size is marked on the left side of the gel with red band.